

Computational investigation of the binding mode of bis(hydroxyphenyl)arenes in 17 β -HSD1: molecular dynamics simulations, MM-PBSA free energy calculations, and molecular electrostatic potential maps

Matthias Negri · Maurizio Recanatini ·
Rolf W. Hartmann

Received: 1 February 2011 / Accepted: 26 July 2011 / Published online: 6 August 2011
© Springer Science+Business Media B.V. 2011

Abstract 17 β -Hydroxysteroid dehydrogenase type 1 (17 β -HSD1) catalyzes the last step of the estrogen biosynthesis, namely the reduction of estrone to the biologically potent estradiol. As such it is a potentially attractive drug target for the treatment of estrogen-dependent diseases like breast cancer and endometriosis. 17 β -HSD1 belongs to the bisubstrate enzymes and exists as an ensemble of conformations. These principally differ in the region of the β F α G'-loop, suggesting a prominent role in substrate and inhibitor binding. Although several classes of potent non-steroidal 17 β -HSD1 inhibitors currently exist, their binding mode is still unclear. We aimed to elucidate the binding mode of bis(hydroxyphenyl)arenes, a highly potent class of 17 β -HSD1 inhibitors, and to rank these compounds correctly with respect to their inhibitory potency, two essential aspects in drug design. Ensemble docking experiments resulted in a steroidal binding mode for the closed enzyme conformations and in an alternative mode for the opened and occluded conformers with the inhibitors placed below the NADPH interacting with it synergically via π - π stacking and H-bond formation. Both binding modes were investigated by MD simulations and

MM-PBSA binding free energy estimations using as representative member for this class compound **1** (50 nM). Notably, only the alternative binding mode proved stable and was energetically more favorable, while when simulated in the steroidal binding mode compound **1** was displaced from the active site. In parallel, ab initio studies of small NADPH-inhibitor complexes were performed, which supported the importance of the synergistic interaction between inhibitors and cofactor.

Keywords 17 β -HSD1 inhibitors · Docking · MD simulations · Binding free energy · MEP · MM-PBSA · DFT

Introduction

Estrogens, in particular the biologically active 17 β -estradiol (E2), display a wide variety of physiological functions via endocrine as well as intracrine, autocrine or paracrine mechanisms. They are essential for the regulation of the female reproductive system and exert a protective function on the cardiovascular system, against osteoporosis and various forms of neurodegenerative diseases [1–3].

On the other hand, estrogens are also known to function as mitogenic factors: the interaction of E2 with the estrogen receptors (ERs) α and β induces an increased expression of transcriptional genes, crucial for the growth and the development of hormone-dependent diseases like breast cancer (BC) [4] and endometriosis [5]. A continuous supply of excess estrogen is essential to facilitate cancer tissue proliferation. However, major parts of the estrogen levels in the plasma are represented by the weakly active estrone (E1) and its sulfate and they are not sufficient to support this progression, suggesting local (intracrine) production of

Electronic supplementary material The online version of this article (doi:10.1007/s10822-011-9464-7) contains supplementary material, which is available to authorized users.

M. Negri · R. W. Hartmann (✉)
Pharmaceutical and Medicinal Chemistry, Saarland University
and Helmholtz Institute for Pharmaceutical Research Saarland
(HIPS), Campus C2.3, 66123 Saarbrücken, Germany
e-mail: rwh@mx.uni-saarland.de
URL: <http://www.PharmMedChem.de>

M. Recanatini
Department of Pharmaceutical Sciences, University of Bologna,
Via Belmeloro, 6, 40126 Bologna, Italy

estrogens to be the key step in the pathological progress [6].

While the ovaries are the main source of systemic estrogen in premenopausal woman, other sites of estrogen biosynthesis (i.e. breast, adipose, bone and brain) spread throughout the body become the major sources of estrogen beyond menopause. These peripheral sites have the capacity to convert C19 steroids (androgens) to C18 steroids (estrogens), but unlike the ovaries they lack the ability to synthesize the former. Hence, this estrogen production is totally dependent on the availability of circulating C19 precursors. Thus, even though the total amount of estrogen synthesized in these peripheral sites may be small, the local tissue concentrations achieved are probably quite high. They will exert significant biologic effect acting at a local tissue level in a paracrine or intracrine fashion [6–9].

In fact, in tumor tissue androstenedione (C19 precursors) and estrone sulfate are transformed by peripheral aromatase and steroid sulfates into E1. Subsequently, the weakly active E1 is reduced to its highly active form E2. This last step of the E2-biosynthesis (Fig. 1) is catalyzed by 17 β -hydroxysteroid dehydrogenase type 1 (17 β -HSD1 or SDR28C1, according to the new nomenclature [10]; E.C. 1.1.1.62), mostly in the peripheral compartments. 17 β -HSD1 belongs to the short-chain dehydrogenase/reductase (SDR)-family and requires NADPH as cofactor for in vivo function [8, 9, 11, 12].

Overexpression of 17 β -HSD1 in tumor cells (i.e. in breast cancer with respect to normal breast tissue) [13–16] as well as in endometriotic lesions [17] and in situ production of E2, i.e. an increased E2/E1 ratio, point out the pivotal role of 17 β -HSD1 in these estrogen-dependent diseases [15, 18–22]. Thus, the selective inhibition of 17 β -HSD1 can be regarded as a promising therapeutic approach in estrogen-dependent diseases [19] as has been the selective inhibition of 5 α -reductase in androgen-dependent diseases [23]. This therapy should present less side effects

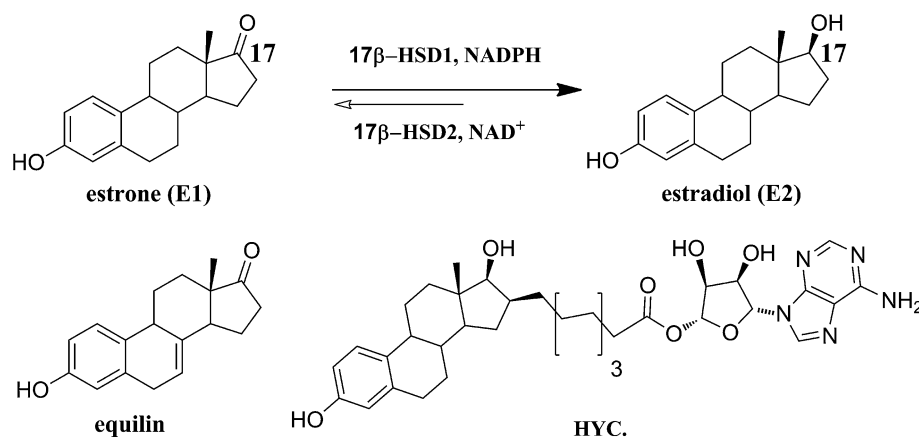
compared to the current endocrine treatments (selective estrogen receptor modulators—SERMs, aromatase inhibitors, GnRH agonists or antagonists, pure antiestrogens) [24–26] in virtue of a selective reduction of the local concentration of E2 in the diseased tissues [27], without influencing the systemic estrogen concentration.

Recently, Husen et al. [11], Lamminen et al. [12] and Day et al. [28] provided a proof of concept for the treatment of breast cancer through the inhibition of 17 β -HSD1 by applying steroidal inhibitors *intraperitoneally* and *subcutaneously* in different mouse-models. In each study a reduction of tumor tissue proliferation was observed.

Several 17 β -HSD1 inhibitors have been reported so far, most of them bearing a steroidal scaffold with substituents at positions 2, 6, 15, and 17 [29–31]. Additionally, five non-steroidal classes exist as well: thienopyrimidinones [32, 33], biphenyl ethanones [34], and from our group, (hydroxyphenyl)naphthalenes [35–37], (hydroxyphenyl)arenes [38, 39], and bis(hydroxyphenyl)arenes [40–43].

Recently, we simulated the bi-bi kinetic cycle of 17 β -HSD1 and performed a structural analysis of the 17 β -HSD1 crystal structures [44], which are present in the PDB database [45] as apoform, as binary complex with cofactor or with E2/androgens/inhibitors, and as ternary complex with cofactor and E2/inhibitors. This study revealed an overall identical tertiary folding for the enzyme with a rigid cofactor binding site (COF) and a tunnel-like, hydrophobic substrate binding site (SUB). The SUB hosts E2, which is stabilized at the polar ends of the SUB by hydrogen bonds between O3 and His221/Glu282, as well as between O17 and Ser142/Tyr155. Notably, Ser142 and Tyr155 are also part of the catalytic tetrad (together with Asn114 and Lys159), essential for the enzyme functionality [44]. The only but marked difference in the crystal structures was identified for the highly flexible β F α G'-loop (residues 188–199), not fully resolved in eleven crystal structures. This loop is situated at the intersection of COF and SUB and can exist in different orientations strongly influencing

Fig. 1 Reduction from E1 to E2 and the steroidal inhibitors equilin and HYC (O5'-[9-(3,17 β -Dihydroxy-1,3,5(10)-estratrien-16 β -yl)-nonanoyl]adenosine)



volume and shape of SUB and COF. Superimposition and successive clustering of the crystals according to the topology of the loop residues Phe192-Met193-Glu194-Lys195-Val196 resulted in three and five clusters, depending on whether the backbone RMSD (root mean square distance) or the all-atom RMSD are considered. The first three clusters correspond to the opened, occluded, and closed states of the enzyme, highlighting the presence of two possible entrance/egress paths for cofactor and ligands (Fig. 2 and ESM1). The last five clusters, on the other hand, discriminate between separated and fused, large and tiny cofactor and substrate binding sites and resemble the main kinetic cycle steps of 17 β -HSD1 as resulted by multiple molecular dynamics (MD) simulations. The outcome of that study suggested 17 β -HSD1 to follow a five-step rapid equilibrium random bi-bi sequential kinetics, in accordance to existing biochemical data. Further, it highlighted the prominent role of the β F α G'-loop in ligand stabilization as well as in entrance/expulsion of cofactor and E1/E2. In particular, Phe192 seems to have two functions: (1) when it points perpendicularly toward Tyr155 it acts as a catalysis promoter by modulating the pK_s of this catalytically crucial residue, thus facilitating the proton transfer to the O17 of E1; (2) it might function as a swinging door, responsible for the opening and the closure

of the entrance/egress gates, as suggested by its marked rotation around the β F α G'-loop axis [44].

Bis(hydroxyphenyl)arenes are highly potent 17 β -HSD1 inhibitors based on a (*m*-hydroxyphenyl)-X-(*p*-hydroxyphenyl) core, where X stands for different (hetero)aromatic rings (Fig. 3) [41, 42]. This class of inhibitors presents a very sharp structure–activity relationship (SAR), where the best substitution pattern is *m*-OH/*p*-OH and the inhibitory potency varies strongly exchanging X. In two recent publications we reported about the two strategies that we followed in the drug design of this class [41, 42]. The first, structure-based, was an ensemble docking approach with the enzyme conformers diverging significantly only in the region of the highly flexible β F α G'-loop [42]. Two strictly conformation-dependent binding modes emerged: a steroid-mimicking one for the enzyme in its closed state (PDB IDs 1fdtB and 1a27) and an alternative one for opened and occluded enzyme structures (PDB IDs 1fdtA and 1i5r). In the latter binding mode, the inhibitors are docked below the cofactor, stabilized by a strong hydrogen bond between the *m*-OH group and the phosphate group close to the nicotinamide ring, as well as by π – π stacking with the nicotinamide moiety of NADPH. The inhibitor is flanked on one side by Tyr155 and Ser142 and on the other by the β F α G'-loop and the α G'-helix. However, no scoring

Fig. 2 Schematic representation of the active site of 17 β -HSD1 and of the putative entrance or/and exit gates 1–3 for the cofactor NADP(H), E1/E2 or the inhibitors

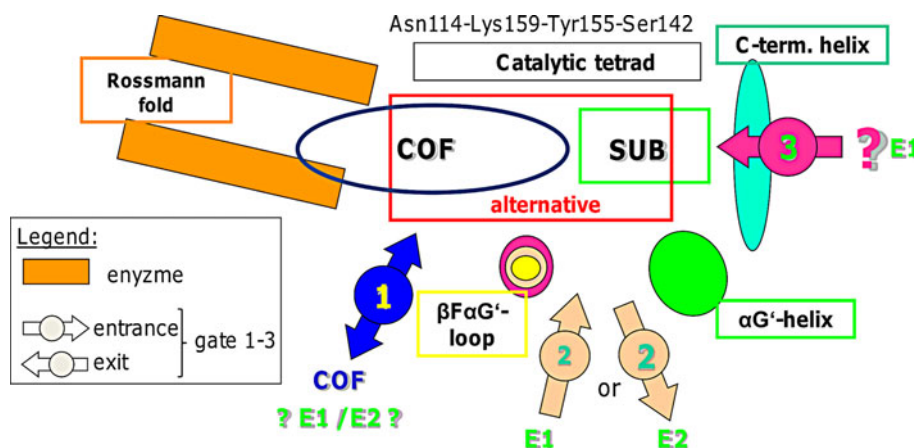
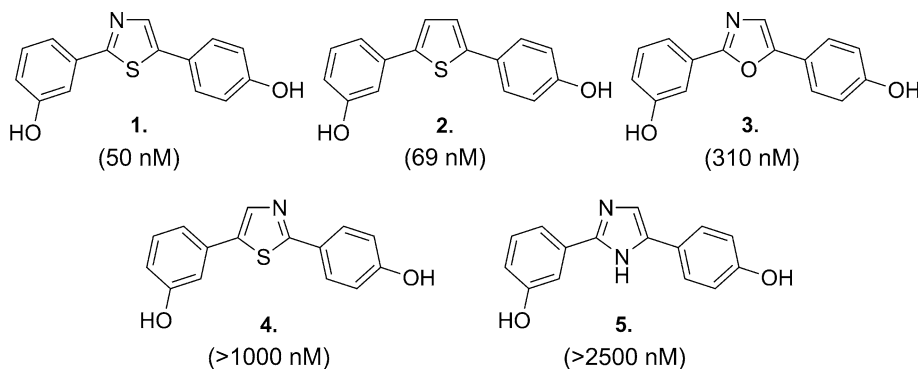


Fig. 3 Structures and IC₅₀ values of compounds 1–5, representative members of the class of the (*m*-hydroxyphenyl)-X-(*p*-hydroxyphenyl)arenes. A strong dependency between the IC₅₀ values and the nature of central heterocycle X has been observed



function managed to rank the docked inhibitors (i.e. 1–5; Fig. 3) correctly according to their IC_{50} values. The second approach was ligand-based and built on the different molecular electrostatic potential (MEP) distributions between active and inactive members of this class [41]. A semi-quantitative MEP-activity relationship (“semi-QMAR”) was derived: electron-rich (negative MEP) *m*-hydroxyphenyl- (region I) and heteroring X (region II), as well as a hydrophobic, neutral *p*-hydroxyphenyl ring (region III), correlate with strong inhibition, whereas switching the MEP distribution (neutral region I and negative region III) results in a drastic loss of activity. In the light of both the alternative binding mode and the characteristic MEP distribution for the ligands we also calculated the MEP distribution for NADPH: the complementarity between the potential surfaces of regions I–II and the nicotinamide moiety supported the preference for this binding mode [42].

Design of the study

In this work we aimed to elucidate the most plausible binding mode for the class of the bis(hydroxyphenyl)arenes, to clarify the role of the flexible $\beta F\alpha G'$ -loop in inhibitor binding, and to find a measure to correctly rank the members of this class of inhibitors with respect to their inhibitory potency. Thereby we grounded our work on the results of the above reported studies, namely on a set of representative enzyme conformers [44] and on the binding modes identified therein [42]. In particular, we focused on the binding mode of compound 1 (Fig. 3), representative for this class of inhibitors.

Again two approaches were followed: in a structure-based perspective, the two different binding modes observed for compound 1 (steroidal and alternative) and the influence of the $\beta F\alpha G'$ -loop on compound 1 binding were studied by means of a multiple-trajectory/-complex MD simulation strategy. Four enzyme conformers, representative for the opened (PDB IDs 1fdtA and 1iol), occluded (PDB ID 1i5r) and closed (PDB ID 1fdtB) enzyme conformers, were used as starting structures with 1 placed either in the alternative (1fdtA, 1iol and 1i5r) or in the steroidal binding mode (1fdtB; two specular orientations) (Fig. 4). Molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) methods [46, 47] and normal mode (NMODE) analysis [48] were employed to approximate the free energies of the five complexes, in order to determine the energetically most favourable binding conformation of compound 1.

The parallel, ligand-based approach consisted in building five ternary complexes based on the catalytic tetrad residues (Asn114, Ser142, Tyr155 and Lys159), NADPH and one of the inhibitors 1–5. We investigated these

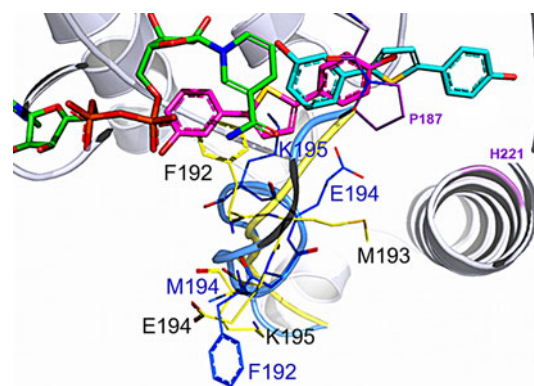


Fig. 4 Docking poses of compound 1 in the closed (1fdtB—yellow; steroidal binding mode; 1 in cyan) and in the opened state of the enzyme (1fdtA—blue; alternative binding mode; 1 in magenta). The cofactor NADPH (green) as well as the $\beta F\alpha G'$ -loop residues F192–K195 are shown in sticks. P187 and H221 are labeled in purple

complexes at density functional theory (DFT) level and the energy differences between these five complexes and the NADPH-tetrad (binary) complex were computed and correlated with the IC_{50} values of the inhibitors. The molecular electrostatic potentials of the ternary complex with 1 and of the binary complex were plotted on the respective electron-densities to compare and the resulting MEP maps were interpreted.

Finally, we docked both active and inactive bis(hydroxyphenyl)arenes into enzyme conformers retained from the MD simulations aiming to rationalize the SAR for this class of inhibitors based on their binding mode.

The combined analysis of both strategies shed new light on the binding mode of this class of inhibitors, on the role of the $\beta F\alpha G'$ -loop, and indicated the alternative binding mode as the most likely one. Compounds 1–5 could finally be ranked correctly based on the relative free energy (ΔG) obtained at a DFT level. We could show the importance of an accurate sampling of the conformational space prior to the selection of the starting structure to be used for structure based-design. Finally, this class of inhibitors seems to follow a novel inhibitory mechanism, namely to inhibit the enzyme dynamics by targeting a specific enzyme conformation and by binding to the cofactor in a synergistic manner instead of competing with it for the cofactor binding site.

Results and discussion

Assessing the binding mode of compound 1 with MD simulations

First, we wanted to investigate the binding mode of compound 1, representative for the bis(hydroxyphenyl)arenes, and to prove the importance of an accurate selection of the

starting enzyme conformation. To this end, and in order to sample most of the conformational space of the flexible β F α G'-loop, five MD simulations lasting at least 10 ns of the ternary enzyme (E)-NADPH-compound **1** complex were performed, where E indicates, respectively: (1) the opened state conformer of 17 β -HSD1 (1fdtA-**D1**)—characterized by a fused cofactor (COF) and substrate binding site (SUB) and by Phe192 turned outwards; (2 and 3) two semi-opened (occluded) conformers (1iol-**D2**; 1i5r-**D3**)—with Phe192 more buried than in 1fdtA and with the loop rotated inwards and toward the COF; (4) the closed state (1fdtB-**D4a** and **D4b**)—mimics the ternary complex at the catalytic moment and presents a closed SUB and Phe192 pointing inwards perpendicular to the catalytic Tyr155.

We expected the most reliable binding mode of compound **1** to be maintained for as long as possible, regardless whether the alternative (**D1–D3**) or the steroidal one (**D4a–4b**) were considered, and eventually to see the five simulations converging into a common final structure. All MD simulations were evaluated in terms of: (a) displacement of compound **1** with respect to the starting pose, (b) complex-stability in terms of RMSD, (c) time required to reach and to maintain a stable RMSD and energy plateau. For each constant segment of the trajectories, snapshots were collected and used to estimate the binding free energy (ΔG_{bind}) by means of MM-PBSA methods and NMODE analysis. This approach of binding free energy determination is less rigorous from a statistical mechanical perspective than thermodynamic integration (TI) or free energy perturbation (FEP), but it is computationally much less demanding and its accuracy has been shown in several recent publications, in particular when ligands with a similar structure were considered [49–51]. In this study ΔG_{bind} were estimated for the very same ligand, namely compound **1**, enforcing us to consider the error in the approximation as negligible.

Another way to reduce the error percentage in (absolute) ΔG_{bind} determination is to calculate the relative binding energy ($\Delta \Delta G_{\text{bind E1-1}}$) between the substrate E1 and compound **1**, for each simulated complex (enzyme conformer) respectively. The impact of the tricky conformational entropy will be reduced and more accurate binding free energy estimations become feasible (Table 1). The binding free energies of E1 (ΔG_{E1}) for the three enzyme-NADPH-E1 complexes with 1fdtA, 1i5r, and 1fdtB were retrieved from the literature [44].

In order to elucidate the dynamics of the active site residues, which modulate the active site topologies and thereby influence the binding of compound **1**, distance collective variables (CV) were derived and plotted over the simulation time. The C_{α} of Met147, which is placed in the substrate binding site and which showed only minor RMSD fluctuations in the MD simulations **D1**, **D3**, **D4a** and **D4b**

(see Figs. 5a, 8a), was selected as reference point for most of the CV measurements. In case of **D2** we chose the C_{α} of Leu149 instead of Met147, because for the latter residues a C_{α} -RMSD of about 1.5 Å was observed. These CV are defined as the distance between the reference point and the center of mass of the phenyl rings of Phe192, His221, Phe226 and Phe259, as well as of the thiomethyl group of Met193 and the amine moiety of Lys195, respectively. All plots of the distance CV over the simulation time can be found in supporting information.

Alternative binding mode

For three ternary complexes describing the alternative binding mode MD simulations of at least 10 ns length were performed: **D1**-1fdtA, **D2**-1iol, and **D3**-1i5r. The starting enzyme conformers of these simulations are characterized by the progressive inward rotation of Phe192, a known “conformational marker”, and by the presence or absence of a short helix in the loop [44]. In each of the three starting conformations compound **1** is close to the cofactor, while the flexible β F α G'-loop is present in divergent conformations.

All three MD simulations **D1–D3** result in a RMSD plateau maintained for about 6 ns (Fig. 5). Nevertheless, the time required to reach these plateaus varies in accordance with the different starting loop conformations. For **D2**, in which the starting complex structure already presents an ordered loop (α -helix: A191-Val196) placed centrally in between the α G-helix and Arg37, a stable all-atom RMSD of 3 Å was reached after 1 ns, while in **D1** (disordered central placed loop) and in **D3** (disordered and left shifted loop) it takes ca. 3 and ca. 6 ns, respectively (Fig. 5b). Interestingly, for all three simulations the final position of the β F α G'-loop axis (C_{α}) is very similar, whereas the orientation of the side chains of the loop residues Phe192 and Met193 diverge significantly (Fig. ESM2—final stable structures of **D1** and **D3** superimposed).

In **D1–D3** compound **1** remains strongly anchored below the cofactor via a hydrogen bond between its *m*-OH group and the phosphate close to the nicotinamide ring and by π - π stacking interactions with the nicotinamide moiety itself (Fig. 6). However, in **D1** two additional hydrogen bonds were formed: one between the *m*-OH group of **1** and the inwardly rotated Lys195 and the second between the *p*-OH group of **1** and His221 (Fig. 7, ESM3, ESM4 and ESM6e).

In **D1** the hydrogen bonds between compound **1** and Lys195 and His221 are formed early in the simulation (Fig. ESM4). This goes ahead with the consistent rearrangement of the loop, which seems to fold into an α -helix. The impact of this motion on the position of **1** is reflected by the RMSD of **1** of nearly 5 Å after few ns, which then sinks to

Table 1 MM-PBSA estimations of binding free energy of compound **1** in the MD simulations **D1–D3** and **D4a–4b**(a) Free binding energies (ΔG_{bind}) for the complexes **D1–D4b** computed by means of MM-PBSA methods and NMODE analysis

E ^a	PDB	MDS ^b	ELE ^c	VDW ^d	GAS ^e	PBSOL ^f	PBTOT (ΔH) ^g		TSTOT (TΔS) ^h		ΔG ⁱ _{bind}
			Mean	Mean	Mean	Mean	Mean	SE	Mean	SE	Mean
Alternative binding mode—enzyme-NADPH-1											
Opened	1fdtA	D1	−36.01	−34.01	−70.02	44.09	−25.94	0.36	−16.75	0.74	−9.19
Occluded	1iol	D2	−18.35	−34.32	−52.67	31.25	−21.41	0.52	−11.94	0.43	−9.48
Occluded	1i5r	D3	−24.49	−36.27	−60.76	36.32	−24.45	0.55	−20.33	0.69	−4.12
Steroidal binding mode—enzyme-NADPH-1											
Closed	1fdtB	D4a	−11.36	−37.49	−48.84	28.88	−19.97	0.49	−17.49	0.78	−2.48
	2 ns extension of D4a		−13.3	−37.18	−50.48	26.42	−24.06	0.41	−13.78	0.56	−10.28
Closed	1fdtB	D4b	−7.6	−30.12	−37.71	21.61	−16.09	0.44	−19.66	0.62	3.66

(b) Relative binding affinity ($\Delta \Delta G_{\text{bind}}$) of compound **1** determined as the difference between ΔG_{bind} and ΔG_{E1} , obtained for the same enzyme conformer, respectively. ΔG_{E1} values taken from Ref. [44]

MDS ^b	ΔG_{E1}^i	$\Delta \Delta G_{\text{bind}} = \Delta G_{\text{bind}} - \Delta G_{\text{E1}}$	$\beta\text{F}\alpha\text{G}'\text{-loop}$	Binding mode
D1	−2.6	−6.59	Open (1fdtA)	Alternative
D3	−0.96	−3.16	Occluded (1i5r)	Alternative
D4a	−8.31	5.83	Close (1fdtB)	Steroidal
D4b	−8.31	11.97	Close (1fdtB)	Steroidal

 ΔH , ΔS and ΔG_{bind} values reported are calculated for the stable RMSD plateau of each MD; (Mean) mean value; (SE) standard error of the mean; all energies expressed in kcal/mol^a (E) starting conformation of the enzyme^b (MDS) molecular dynamics simulation code^c (ELEC) electrostatic contribution in gas phase^d (VDW) Van der Waals contribution in gas phase^e (GAS) free energy in vacuum^f (PBSOL) solvation energy^g (PBTOT) (ΔH) enthalpy^h (ΔS) (TSTOT) entropic contributionⁱ (ΔG_{bind}) free binding energy^j (ΔG_{E1}) free binding energy calculated for the substrate E1 by means of MMPBSA/NMODE based on MDS with the different enzyme conformers

4 Å after 5 ns and is maintained until the end of the MD (Fig. 5c). The main responsible for this large fluctuation is the *p*-hydroxyphenyl-ring, which is initially pushed toward the active site wall where Met147 is placed (4 Å at 2 ns) and then close to the α G-helix (8 Å plateau after 3 ns; Fig. ESM6f). His221 turns slightly inward, breaking the salt bridge formed with Glu282 (Fig. ESM5a), and establishes in turn a hydrogen bond with the *p*-OH group of compound **1** (0% occupancy at 1 ns; 79% from 3 to 12 ns; Fig. ESM4). As a direct consequence of these movements a concerted motion between the C-terminal helix and neighboring structural motifs can be observed. The C-terminal helix slides up for about 20° and brings the carboxylic moiety of Glu282 within 3–4 Å of the guanidine moiety of Arg258 with which it forms a new ion-pair (Fig. ESM5b and ESM6c). This Arg258-Glu282 salt bridge

occludes the access to the substrate binding site at the C-terminal end (gate 3, Figs. 2, 6a). Remarkably, in the PDB ID 1equ Arg258 contributes to stabilize the steroidal inhibitor equilin. Phe226 moves inward (from 10 to 4 Å after 4 ns) and after 6 ns it stabilizes the central thiazole ring of **1** by π – π stacking (Fig. ESM6d). Phe192 and Met193 are still placed out of the active site, with the former shifted toward the adenosine moiety of NADPH, thus hinting at a partial closure of the entrance cleft (gate 1, Figs. 2, 6a) of the cofactor.

In the stable conformation of **D2** the $\beta\text{F}\alpha\text{G}'\text{-loop}$ is folded into a long α -helix (Ala191-Leu197). This helix already exists in the starting enzyme conformation and the initial orientation of the loop residues might reduce the conformational freedom of the loop itself. In the first, very stable 6 ns of **D2** Phe192 points toward the cofactor,

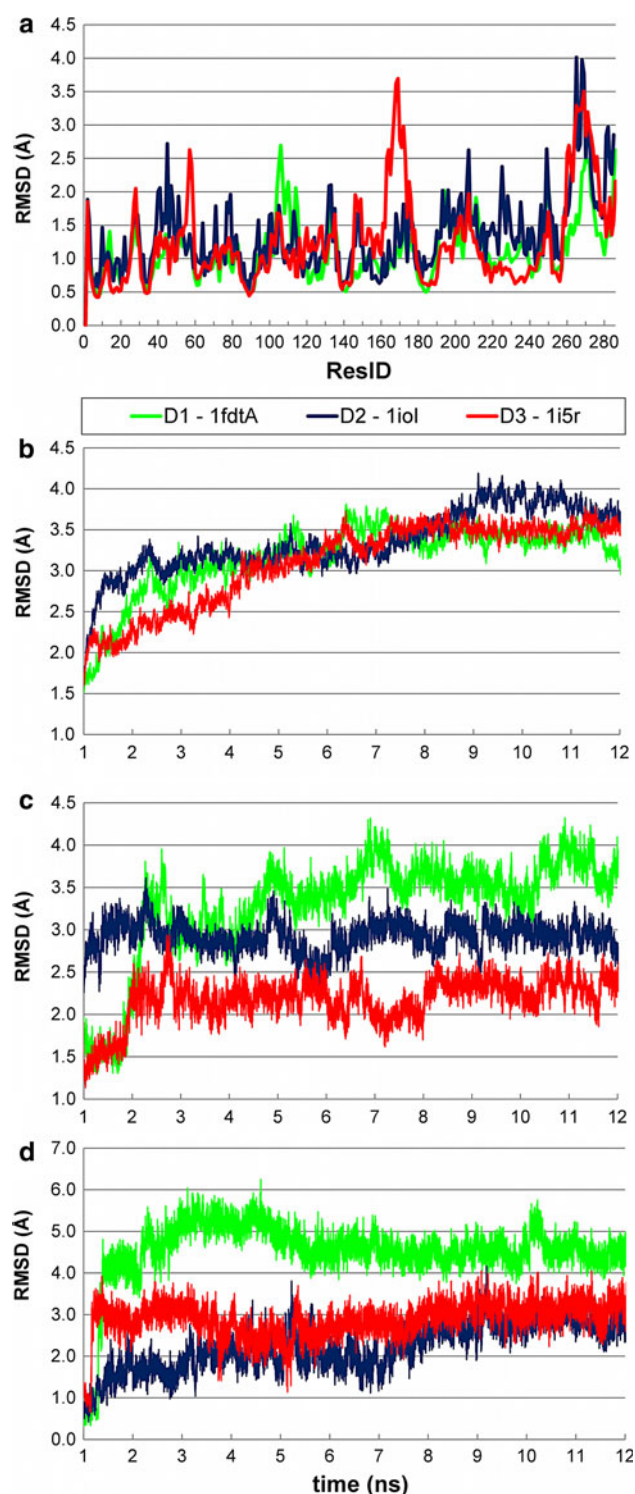


Fig. 5 Trajectory analysis of the MD simulations **D1** (green), **D2** (red) and **D3** (blue) (alternative binding mode) reported in terms of **a** residue-dependent RMSD fluctuation, **b** time-dependent all-atom RMSD fluctuation for all 285 residues, **c** time-dependent all-atom RMSD fluctuation for the $\beta\alpha\alpha'$ -loop, and **d** time-dependent fluctuation of the inhibitor with respect to its starting (docking) pose

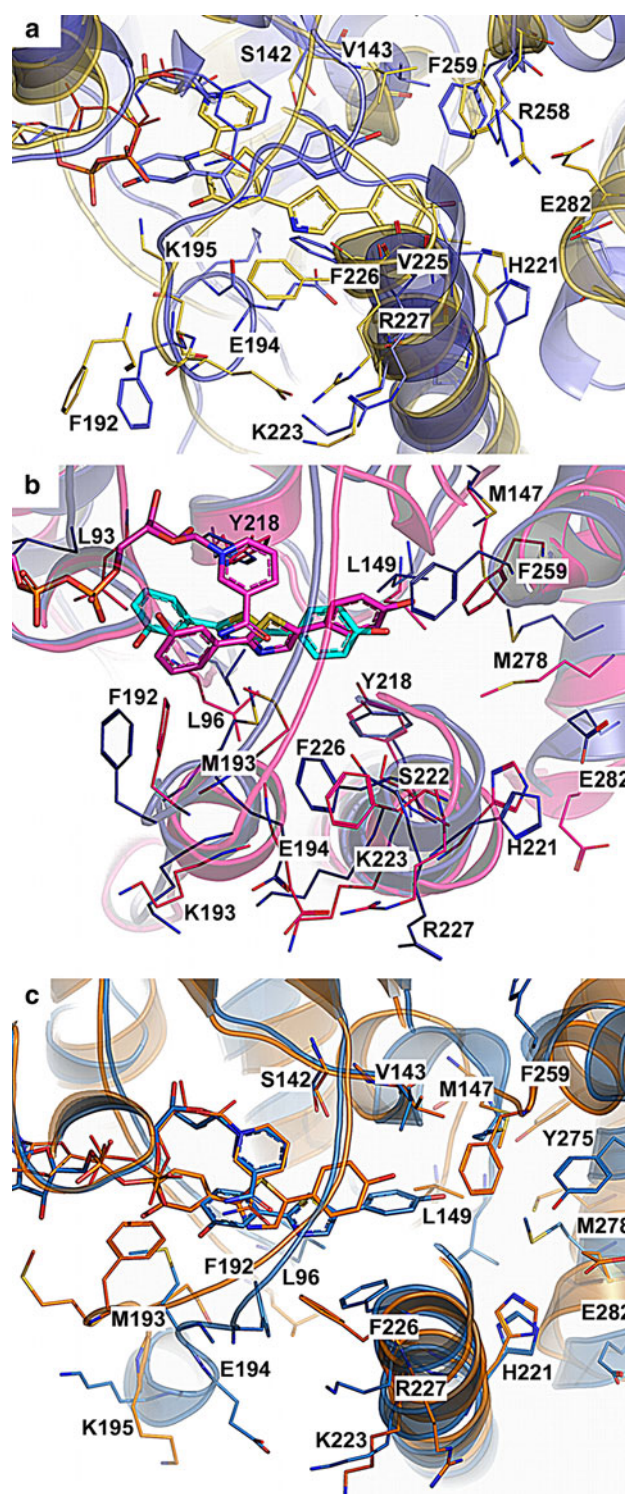


Fig. 6 Analysis of the MD simulations **D1**, **D2** and **D3** with compound **1** in the alternative binding mode. **a** Overlay of the initial (blue) and the average structure of the stable segment (yellow) of **D1**. **b** Overlay of the initial (blue/cyan) and the average structure of the stable segment (magenta) of **D2**. **c** Overlay of the initial (orange) and the average structure of the stable segment (blue) of **D3**. Important residues are rendered as lines, compound **1** and NADPH as sticks

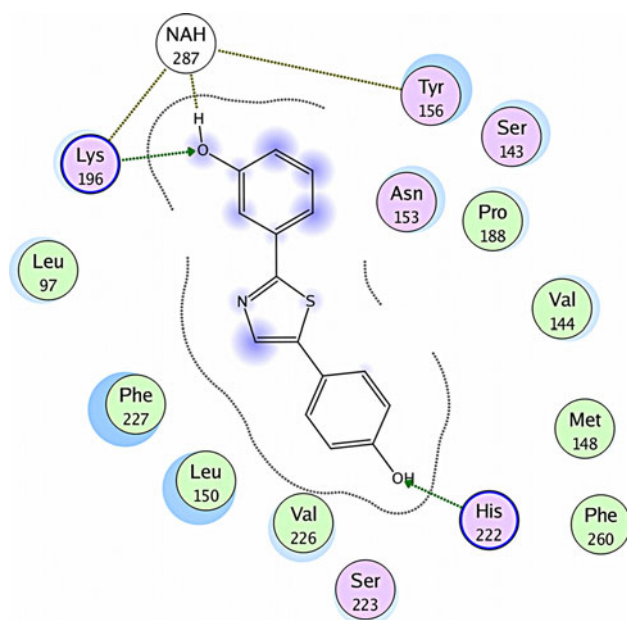


Fig. 7 Schematic representation of the binding mode of compound **1** in the stable part of **D1**. Apolar residues are colored in *green*, while hydrophilic ones in *magenta*. The H-bonds between **1** and Lys195, H221 and the cofactor NADPH (here NAH) are highlighted by dotted lines

closing gate1, and it moves close to compound **1** interacting via π -stacking with the *m*-hydroxyphenyl ring (Fig. 6b). Met193 is buried into the active site playing an important role in the stabilization of the inhibitor (Fig. ESM7b and d) and it induces the central ring X of the inhibitor to be perpendicular with respect to the nicotinamide moiety of NADPH and parallel to the catalytic Tyr155 (parallel π - π stacking). After 7 ns a right-shift of the loop axis toward the α G-helix was observed, causing an inward rotation of the phenyl ring of Phe226 of nearly 3 Å (Fig. ESM7c). This motion closes gate 2, placed between the β F α G'-loop and α G-helix (Fig. 6b), but opens gate 1 and induces a consequent drop of the cofactor. Thereby the hydrogen bond between the amide moiety of the nicotinamide ring of NADPH and the backbone of Val188, characteristic of cofactor stabilization in several HSD enzymes, is lost (72% occupancy 0–6 ns, 24% 6–12 ns; Fig. ESM4 and ESM8—stable (6 ns) and final pose of **1** in **D2**) [52]. Notably, the inward rotated Phe192 holds NADPH in the cofactor binding site and the hydrogen bond of the *m*-hydroxy group of compound **1** to the phosphate of the cofactor is maintained.

In the stable part of **D3** compound **1** interacts via π - π stacking with Phe192, this time facing the active site, and with Phe226, which stabilizes the *p*-hydroxyphenyl ring (Fig. 6c and ESM9). Additionally, **1** forms hydrophobic interactions with Leu96, Val143, Leu149, and Val225. In this case Met193 is oriented toward to adenosine moiety of

NADPH shielding the cofactor binding site from the solute like Phe192 does in **D2**. The *p*-hydroxyphenyl ring (reg III) of **1** points into a pocket formed by Met147, Tyr275 and Met278. Already after few ns the β F α G'-loop is folded as α -helix (Met193 to Leu197), apparently avoiding the concerted, disordered motions of C-terminal helix and α G-helix residues observed for **D1**. The ion-pair Glu282-His221 is maintained (Fig. ESM5c and ESM9d) and the C-terminal helix does diverge marginally with respect to the α G-helix and from its starting structure (lower RMSD fluctuations of residues 275–285 of **D3** with respect to **D1** in Fig. 5a).

In all three simulations, Glu194 is involved in ion-pairs with the α G-helix residues Lys223 and/or Arg227 (Fig. ESM5), stabilizing the central position of the loop and modulating gate 2. However, depending on how close loop and α G-helix are, in **D1** and **D3** gate 2 is reduced to a tiny hydrophobic tunnel, delimited by Val196, Leu197, Gly198 (β F α G'-loop) and Gln217, Ala220 and Phe226 (α G-helix), while in **D2** it is totally closed. Remarkably, Lys195 is rotated toward the adenosine moiety of NADPH only in **D1** and **D3**, forming a salt-bridge with the 2' phosphate of the cofactor, while for **D2** it is solvent exposed and further afar from the phosphates, thus leaving gate 1 opened and exposing NADPH to the solvent.

Overall, the final scenario emerging from **D1–D3** suggests an occluded active site similar to the one of crystal structures resembling the catalytic moment (e.g. PDB ID 1a27). In **D1** and **D3**, but also in the first 6 ns of **D2**, the β F α G'-loop perches in between cofactor and substrate binding site, stabilized in its central position by interactions with the α G-helix (ion pair Glu194-Lys223/Arg227) and the 2'-phosphate of the cofactor (Fig. 6). It occludes both gate 1 and 2, avoiding the solvent to enter the substrate binding site and preserving its hydrophobic nature. Further, in this position the β F α G'-loop hampers the release of cofactor and inhibitor.

The ΔG values (Table 1a) for **D1**, **D2** (first 6 ns), and **D3** are −9.19, −9.48, and −4.12 kcal/mol respectively and appear to support our conclusions. The ΔG of the complexes **D1** and **D2** are in line with the high potency of **1** (50 nM), probably because of the presence of the helix in the β F α G'-loop, which can stabilize the inhibitor-enzyme complex in a local minimum energy conformation. Interestingly, also **D3** presents a helix in the loop but its estimated ΔG is much lower. The reason therefore is reflected by the higher entropic term obtained for **D3**, whereas its enthalpy is comparable to that of **D1** (Table 1a).

Steroidal binding mode

The two MD simulations with compound **1** placed in the substrate binding side (SUB), **D4a** (*m*-OH mimicking the

C17 OH and the *p*-OH the C3 OH of estrone) and **D4b** (**1** is horizontally flipped with respect to **D4a**), reach a longer all-atom RMSD plateau only after 11 ns (Fig. 8b). However, several shorter plateaus of approximately 1 ns occur along the trajectories as well, indicative for possible intermediate metastable states as also suggested by the exchange of hydrogen bond partners of compound **1** along the simulation time (see % occupancy in Fig. ESM4).

In **D4a** the β F α G'-loop shows an all-atom RMSD of about 4 Å after already 1 ns, which sinks to 3 Å after 2–3 ns before rising again to 4 Å over the last 5 ns (Fig. 8c). The rapid RMSD increase at the beginning of the trajectory is caused by the shift of the loop toward the α G-helix, protruding into the SUB with Phe192 and Met193, as well as Phe226, turning more and more into the active site (Fig. 9): Thereby, Phe192 moves from nearly 20 Å to 11 Å from the C $_{\alpha}$ of Met147, whereas Phe226 from 15 to 7 Å (Fig. ESM10). The initial motion of the loop induces the displacement of the inhibitor (heavy atom RMSD shift of more than 5 Å; Fig. 8d) and thereby the loss of the hydrogen bond between compound **1** and Ser142/Tyr155 as well as His221/Glu282 (0% occupancy after 3 ns and 1 ns respectively; Fig. ESM4 and ESM5d-e). The *m*-hydroxyphenyl-ring of **1** is progressively oriented towards the bottom of the SUB, where it forms a hydrogen bond with Asn152 (% occupancy rises from 0 to 80%; Fig. ESM4), whereas the *p*-hydroxyphenyl-ring moves towards the top of the SUB and gate 3, inducing Phe259 to rotate outwards, as shown in Fig. 9 and ESM10.

In **D4b** a similar behavior of the β F α G'-loop as for **D4a** was observed with an RMSD of about 3.5–4 Å after 2 ns (Fig. 8c). Again the loop moves into the SUB and pushes compound **1** progressively toward gate 3. Notably, in this simulation the *m*-hydroxyphenyl-ring of **1** remains anchored to His221 (average occupancy about 19%; Fig. ESM4) following His221 out of the SUB (Fig. ESM11). The *p*-hydroxyphenyl-ring, on contrary, loses its hydrogen bond to Ser142/Tyr155 after approximately 6.5 ns (Fig. ESM4) and is then pulled to the top and out of the SUB (Fig. 9). The jump in the all-atom RMSD of **1** shown in Fig. 8d from 4 Å to a plateau of 8–10 Å summarizes this motion.

For both simulations **D4a** and **D4b** compound **1** is displaced from its starting pose ending either perpendicular to the estradiol plane close to gate 3 (**D4a**) or lying completely out of the SUB (**D4b**) (Fig. 9). The large fluctuations in its RMSD are in agreement with the unfavorable ΔG values estimated for **D4a** (−2.41 kcal/mol) and especially **D4b** (+3.57 kcal/mol) (Table 1). Interestingly, in contrast to **D1–D3**, at the end of **D4a** and **D4b** cofactor and substrate binding site result clearly separated by the

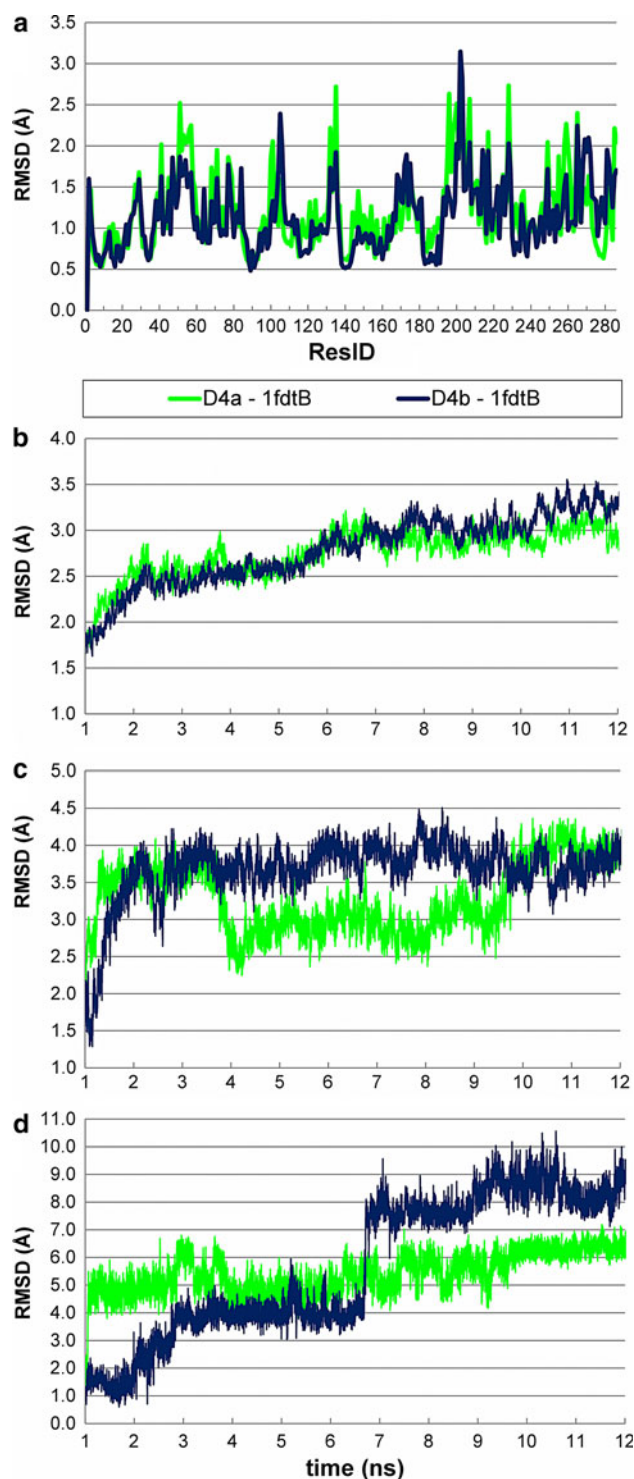


Fig. 8 Trajectory analysis of the MD simulations of the complexes **D4a** (green) and **D4b** (blue) (steroidal binding mode) reported in terms of: **a** residue-dependent RMSD fluctuation; **b** time-dependent all-atom RMSD fluctuation for all 285 residues; **c** time-dependent all-atom RMSD fluctuation for the β F α G'-loop; and **d** time-dependent fluctuation of the inhibitor with respect to its starting (docking) pose

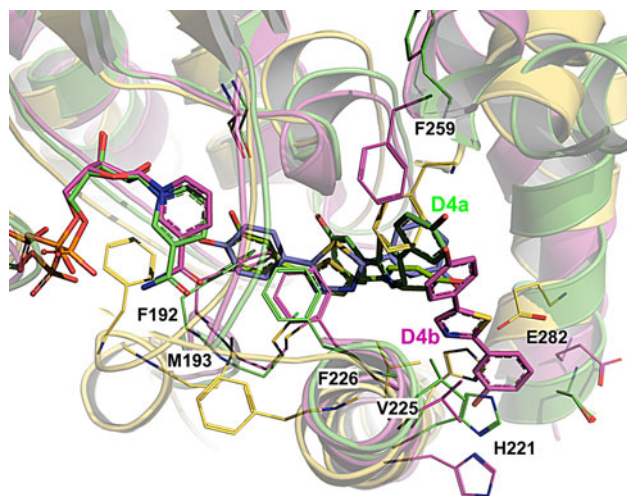


Fig. 9 Analysis of the MD simulations **D4a** and **D4b** with compound **1** in the steroid binding mode. Both the initial (**D4a** light green; **D4b** violet) and the enzyme 3D-structures representative for the stable segment of **D4a** (green) and **D4b** (magenta) are superimposed with the common starting conformation (residues colored in yellow)

β F α G'-loop and Phe192 is buried in the SUB. In **D4b** the salt-bridge His221-Glu282 is broken (Fig. ESM5e) due to the motions of loop and compound **1** and both residues edge outward, enlarging gate 3 (Fig. 9).

In order to minimize error contributions to the binding free energies given by the entropy factor we calculated the relative binding free energies for compound **1** with respect to E1 ($\Delta\Delta G_{\text{bind E1-1}}$; Table 1b). Thereby we subtracted the binding free energies for E1 (ΔG_{E1}), taken from Ref. [44], from the ΔG_{bind} estimated in this study for compound **1**. Both the ΔG_{E1} and the ΔG_{bind} were obtained for the same starting enzyme conformers, i.e. 1fdtA, 1i5r and 1fdtB. The $\Delta\Delta G_{\text{bind E1-1}}$ resulting for **D1** (−6.59 kcal/mol) and **D3** (−3.16 kcal/mol) favour the alternative binding mode in particular comparing them to the $\Delta\Delta G_{\text{bind E1-1}}$ obtained for **D4a** (+5.83 kcal/mol) and **D4b** (+11.97 kcal/mol). Remarkably, the $\Delta\Delta G_{\text{bind E1-1}}$ for **D1** and **D3**, as well as the difference between them, are in concord with the ΔG_{bind} obtained in this study for **D1** (−9.19 kcal/mol), **D2** (−9.48 kcal/mol) and **D3** (−4.12 kcal/mol).

Recently, Mobley et al. showed that when an enzyme is kinetically trapped in a metastable state errors in a range of several kcal/mol might result for binding energy determination and that a solution might be the inclusion of the relative free energy of reorganizing the protein in addition to the binding free energy of the ligand [53]. However, in our study consistent conformational rearrangements of the β F α G-loop were observed in all MD simulations (see the all-atom RMSD plots in Figs. 5b, c, 8b, c), which lead us to exclude the enzyme to be kinetically trapped in a metastable state endorsing to neglect this additional

contribution. Moreover, the fact that **D1–D3** resulted in a common, energetically favorable conformation (focusing on the β F α G'-loop), regardless that different enzyme conformers were employed as starting point to test the same binding hypothesis, accounts for a selected-fit mechanism for this class of inhibitors. In conclusion, although the five MD simulations **D1–D4b** did not converge to one common binding mode, at least for **D1–D3** a similar final pose for compound **1** was found. The β F α G'-loop axis is in the same position for all of the three and similar binding free energies were obtained for the simulated complexes in **D1** and **D2**. Taken together, the common binding pose of compound **1** in **D1–D3**, its large motions and almost “out-of-substrate-binding-site” position in **D4a–4b**, as well as the related ΔG_{bind} values, indicate the alternative binding mode of **1** to be the more plausible one. Moreover, we suggest that **1** would ideally bind to NADPH, Phe192, Phe226 and the inward turned His221, with the C-terminal helix switched up occluding the substrate binding site. The ΔG values and the long, unstable trajectories of **D4a–4b**, which are characterized by various RMSD plateaus and short-living hydrogen bonds with different residues, lead to the conclusion that the steroidal binding mode is less likely for this class of inhibitors and that it might eventually represent a transition step on the egress pathway. This last assumption is supported by the ΔG estimated for an extension of 2 ns of simulation **D4a** (−10.28 kcal/mol, with respect to −2.48 kcal/mol of the stable plateau; Table 1a), where compound **1** already points out of gate 3 and the SUB is reduced by half and occupied by Phe192, Met193, and Phe226. This gain in free energy could be indicative for an ideal egress way through gate 3.

Molecular electrostatic potential calculation and analysis

Molecular electrostatic potential (MEP) typically reveals parts of the molecules more likely to interact, and considering that the interaction energy is usually dominated by the electrostatic part we expect the electrostatic potential to be a sensitive tool for prediction of the potency of inhibitors **1–5**. Thus, a complementarity between the nicotinamide ring of the cofactor and the *m*-hydroxyphenyl ring (region I) of these compounds has been hypothesized on the basis of the opposite molecular electrostatic potential (MEP) distributions for these areas, suggesting π – π stacking between cofactor and inhibitor (Fig. 10) [42].

In order to verify this assumption and the electrostatic nature of the binding, we created a set of five complexes **T1–T5** formed by the catalytic residues (Asn114, Ser142, Tyr155 and Lys159), NADPH and compounds **1–5** (Fig. 3), respectively, and one “receptor” complex **R1** consisting of the catalytic residues and NADPH. All six

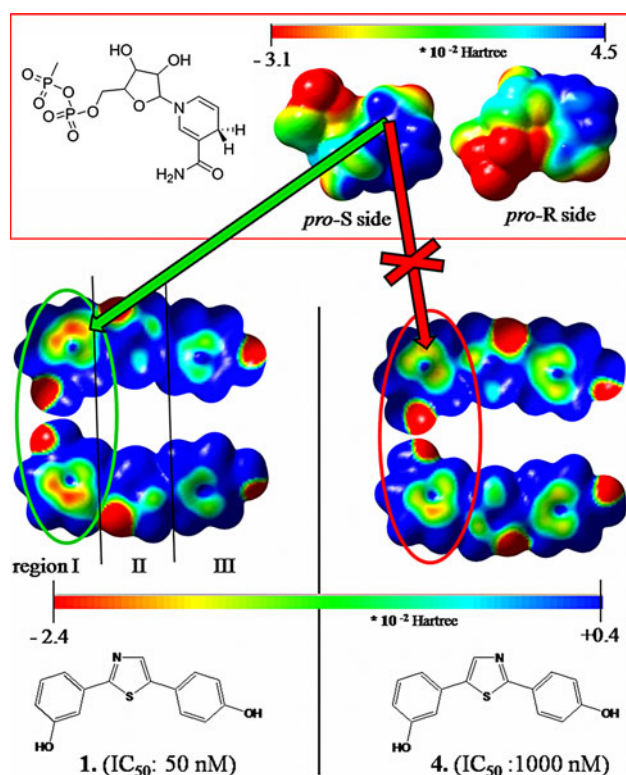


Fig. 10 Differences in the MEP maps of compounds **1** and **4** in region I (*m*-hydroxyphenyl ring) and III (*p*-hydroxyphenyl ring) and the complementarity to the nicotinamide moiety of NADPH

complexes were ab initio geometry optimized at the B3LYP/6-311+g** level of DFT [54–56]. Each of compound **1–5** was placed in the alternative binding mode, anchored with the *m*-OH group to the phosphate next to the nicotinamide moiety, and the (*m*-hydroxyphenyl)-X core parallel to it (Fig. 11). The IC₅₀ values of compounds **1–5** span a range between 50 and 2,500 nM and show a strong dependency on the nature of the heteroring X. The energy differences (ΔG) between each of the five ternary complexes and the binary complex were computed and plotted against the pIC₅₀ values (Fig. 11).

This ab initio approach was not intended to predict the absolute free energy of binding in quantitative agreement with the experimental data, i.e. IC₅₀ values, but rather to determine relative binding free energies to be used to rank correctly the inhibition profile for the structurally very similar inhibitors **1–5**. Remarkably, the calculated ΔG reproduce the trend of **1–5** in terms of inhibitory potency (Fig. 11b), and could be used to correct rank the inhibitors. However, no linear correlation between the pIC₅₀ and the ΔG could be derived. On contrary, the inhibitors rather fall into two classes with compounds **1** and **2** showing a ΔG of approximately $-1,175$ kcal/mol and compounds **3**, **4** and **5** a ΔG around -850 kcal/mol. This last fact necessitates of additional investigations which were beyond the scope of

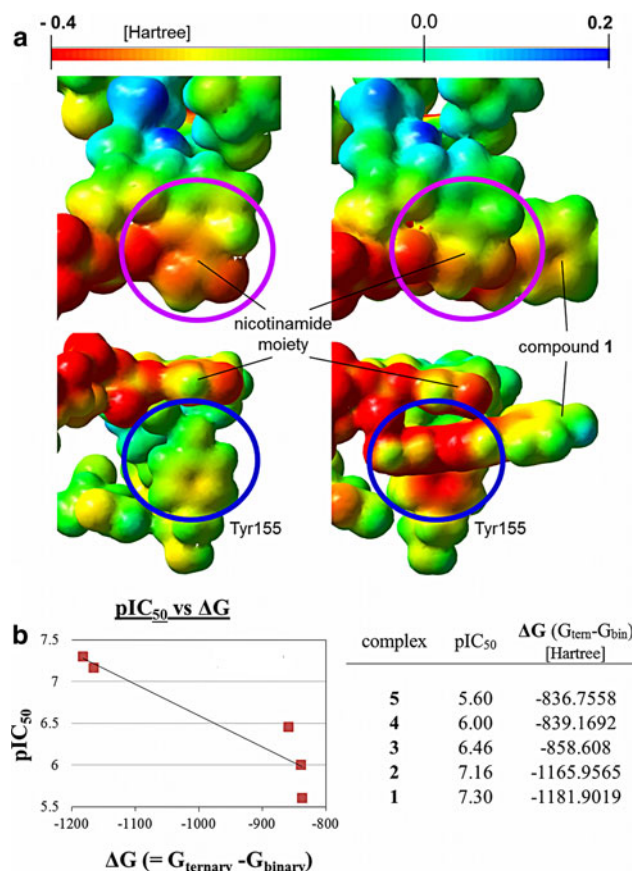


Fig. 11 **a** Differences of the electrostatic potential distribution for binary and ternary complex, enzyme-NADPH-**1** and enzyme-NADPH, respectively, and **b** correlation between the pIC₅₀ values of compounds **1–5** and the energy difference between the two complexes (ΔG)

this study. Nevertheless, two interesting aspects that might explain this clear-cut separation have been considered. The first is the sulfur atom in the central heterocycle X of **1** and **2**, capable of electrostatic interactions with Tyr155. The second, more relevant aspect is that for region I of **1** and **2** the MEP distribution is more negative compared to that of **3**, **4** and **5** [42], a valuable hint for a strong interaction with the nicotinamide moiety. A paradigmatic example is given by compounds **1** (IC₅₀ 50 nM) and **4** (IC₅₀ 1,000 nM), which differ only in the position of the nitrogen on the thiazole ring, and where the MEP distributions for region I and III are inverted (Fig. 10). The difference between their ΔG is consistent, sustaining the importance of specific MEP ranges in region I necessary for the π - π interaction with the nicotinamide moiety.

Thus, we were able for the first time to rank this inhibitor class according to a parameter, ΔG , whereas all conventional scoring methods failed in evaluating the influence of the heterocycle X. Further we could show that a close relationship between MEP distribution of the inhibitors and the interaction with the cofactor must exist.

The representation of the MEP on a molecular isosurface can be a useful method for rationalizing the interactions between molecules or the molecular recognition processes [57]. Thus, the MEP maps for the “receptor” complex **R1** and the “receptor-compound **1**” complex **T1** were calculated by plotting the electrostatic potential on the electron density (Fig. 11). Clear differences emerged for the nicotinamide moiety and the Tyr155 areas of the two maps. The former, in particular, is red for complex **R1** (negative MEP), suggesting a polarized and reactive area, whereas for complex **T1** the nicotinamide ring was shifted to green/yellow in the color scale, indicating a neutral, less reactive area (Fig. 11). This might be due the electron rearrangement occurring for the nicotinamide ring when NADPH and compound **1** interact: the π -electrons from the upper side of the nicotinamide (red in **R1**) are drawn toward the inhibitor where they are involved in the π - π stacking interaction.

SAR elucidation in the light of the alternative binding mode

In order to rationalize the stringent SAR observed for this class of inhibitors of 17 β -HSD1, as well as to verify the final active site conformations emerged from the MD simulations, we docked selected active and inactive derivatives (Table 2) into the stable conformation of **D1**, **D2**, and **D3** using GOLD v4.0.1 [58, 59]. These compounds

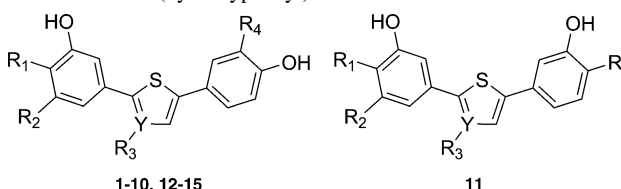
are variously substituted in para (R_1) or meta (R_2) position on the *m*-hydroxyphenyl-ring, on the central arene X (R_3) as well as in meta position on the *p*-hydroxyphenyl-ring (R_4). Remarkably, whenever a substituent caused a modification of the “optimal” MEP distribution ranges, it always resulted in a detrimental effect on the activity.

The most critical substituent seems to be R_1 : already small substituents like Me (**6**), OH (**7**) and F (**8**) reduce drastically the inhibitory potency, most likely hindering the hydrogen bond formation between the *m*-OH group and the phosphate oxygen of NADPH due to steric effects or electrostatic repulsion (Fig. 12a).

In R_2 a Me (**9**) reduces the potency, whereas the electronegative F (**10**) increases it. The reason seems to be the close vicinity to the β FxG'-loop and to Leu96, on one side, as well as to NADPH and the catalytic Tyr155 on the other side. The docking results suggest that already a small methyl group (Me) might hinder the formation of the H-bond to the cofactor (**9**; Fig. 12b), whereas a F could be involved in halogen bonds with NH- of the coil backbone or with the OH of Tyr155 itself.

When compounds with bulky R_3 substituents (i.e. *m*-OH-phenyl-; **11**; Fig. 12d) were docked a different binding pose with respect to that of compound **1** was observed. The characteristic *m*-OH/*p*-OH-motif of the bis(hydroxyphenyl)arenes does not fit anymore, whereas the *m*-OH/*m*-OH-pattern now allows a better placement of the huge R_3 substituent in the above mentioned pocket. This is also

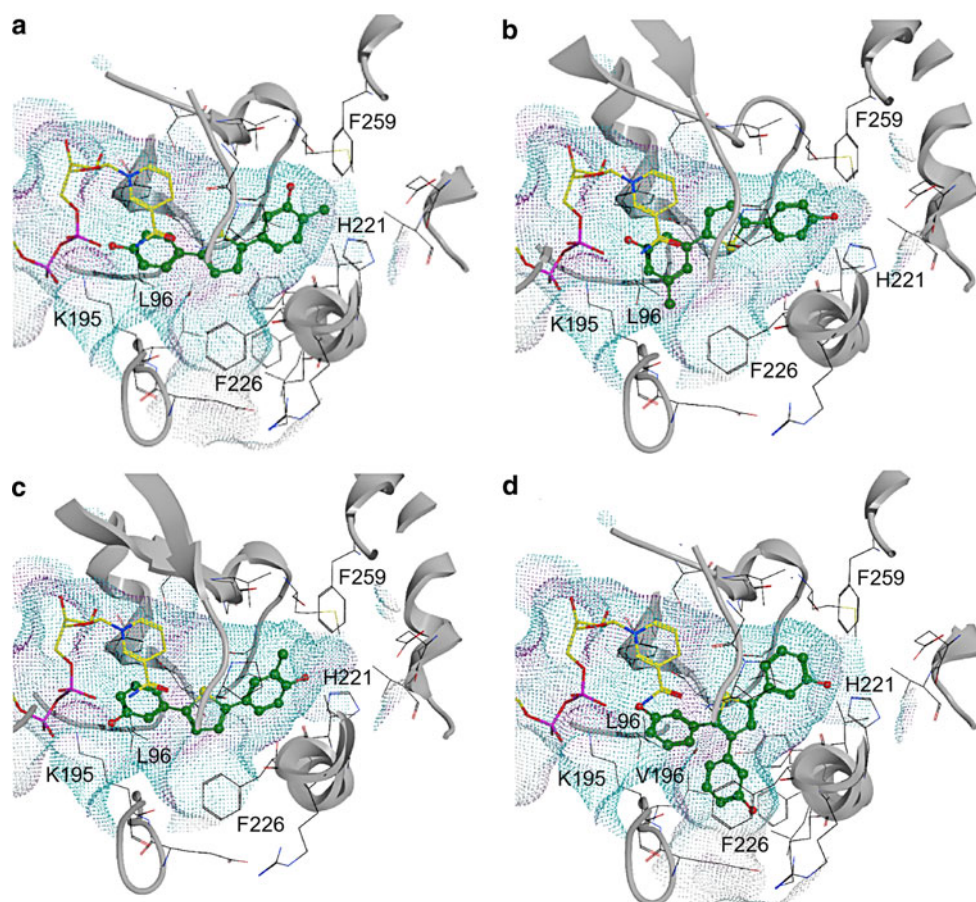
Table 2 17 β -HSD1 inhibitors of the series of the bis(hydroxyphenyl)arenes

						
Compound	R_1	R_2	R_3	R_4	Y	IC ₅₀ (nM) (% at 1 μ M) ^a
1	H	H	H	H	N	50
2	H	H	H	H	C	69
6	Met	H	H	H	C	(19)
7	OH	H	H	H	C	(16)
8	F	H	H	H	C	113
9	H	Met	H	H	C	629
10	H	F	H	H	C	42
11	H	H	<i>m</i> -OH-Ph-	H	C	119
12	H	H	<i>m</i> -OH-Ph-	H	C	(31)
13	H	H	H	F	C	8
14	H	H	H	Met	C	46
15	H	H	H	C ₂ H ₄ COOEt	C	135

Activity differences are observed due to single substitutions on the three ring systems

^a Experimental IC₅₀ and % inhibition values are retrieved from references [40, 41]

Fig. 12 Docking poses of compounds **6** (a), **9** (b), **14** (c) and **11** (d) in the optimized average structure of the stable RMSD plateau of **D1**. The binding pocket is represented in dots, which are color-coded from cyan (hydrophobic) over green (slightly polar) to magenta (polar area). The subpocket formed over the simulation time is delimited by L96 and F226 and it can host an additional *m*-OH-phenyl moiety as seen for compound **11** (d)



substantiated by the IC_{50} values of compounds **11** (119 nM) and **12** ($19 \pm 10\%$ at 1 μ M).

In R_4 various substituents are tolerated, eventually also leading to a marked increase in potency (i.e. F, **13**; Me, **14**). R_4 is close to a hydrophobic pocket (more or less present in all three MD simulations **D1–D3**), which is delimited by Val 143, Met147, Phe259, Leu262 and Met279. R_4 seems to have two functions: on one hand it can modulate the molecular electrostatic potential distribution of the inhibitor, as seen for F in compound **13** [42], and on the other it can directly undergo hydrophobic interactions with the pocket residues (Met, **14**) (Fig. 12c).

Summarizing, we used snapshots from the stable segments of the trajectories of **D1** and **D3** to dock the class of bis(hydroxyphenyl)arenes. Thereby we succeeded to discriminate between active and inactive compounds based on their binding pose and to explain the influence of substituents (e.g. F as R_1) on the inhibitory potency. The existence of the hydrophobic pocket between the lower $\beta F\alpha G'$ -loop region and the N-terminal part of the $\alpha G'$ -helix and its capability to accommodate larger R_3 substituents, limited to **D1** and **D3**, are important findings for future drug design. In particular, it suggests that additional hydroxy-containing moieties might be suitable substituents in this

position, even though this would imply to change the *m*-OH/*p*-OH pattern. At this respect, also an elongation of the inhibitors on the *p*-hydroxyphenyl side, e.g. by exchanging it with a naphthalene core, might be a suitable variation. Finally, the importance of conformational sampling and successive choice of the correct enzyme conformation to be used for drug design was evidenced.

Conclusions

Molecular dynamics simulations and MM-PBSA free energy calculations were applied to infer the binding mode of 17 β -HSD1 inhibitors belonging to the class of bis(hydroxyphenyl) arenes. In three out of five MD simulations compound **1** oriented itself in the alternative binding mode, where it synergistically interacted with the nicotinamide moiety of NADPH. Besides supporting the alternative binding mode as the likeliest for compound **1**, the results of the MD simulations shed new light on the role of the $\beta F\alpha G'$ -loop in the binding of the class of bis(hydroxyphenyl)arenes. The importance of choosing the correct enzyme conformation to be used for drug design was evidenced, since different, conformation-dependent

binding modes could hamper the correct interpretation of the biological results, the SAR and the future drug design.

A global picture emerged from the concerted analysis of the MD studies and the quantum-chemical investigations of the molecular electrostatic potential of five catalytic residues-NADPH-inhibitor complexes. It suggested for this class of inhibitors a novel mechanism of action with respect to that observed for steroidal inhibitors. In fact, the bis(hydroxyphenyl)arenes seem to inhibit the enzyme dynamics by simultaneously binding to active site and cofactor, strongly influenced by electrostatic interactions, and, thus, by “freezing” the loop in a particular conformation. Hence, the enzyme would be restricted into a “half-switching” state, not capable to evolve in its kinetic cycle, in a similar manner as shown for the *E. coli* DHFR inhibitor methotrexate [60].

Materials and methods

Computational details

Crystal structures of 17 β -HSD1 were obtained from the Protein Databank (PDB, www.pdb.org) [45]. In detail, we used the following PDB structures: 1fdt, 1i5r and 1iol. 1fdt presents Arg37 and the loop residues Thr190 to Gly198 in two very different conformations, denoted in this article as 1fdtA (open state) and 1fdtB (closed state). For 1i5r the cofactor NADPH was merged into the enzyme after an accurate overlay with 1fdt, thus the hydride inhibitor HYC (1i5r) was removed and NADP⁺ was turned into NADPH. Water molecules were removed and correct atom types set by means of the software package MOE 2008.10 [61]. Ionization states and hydrogen positions were assigned using the Protonate 3D utility of MOE, histidines oriented toward the outer part of the enzyme were considered as protonated after a prediction run made by MolProbity [62, 63]. Thus, the enzymes were energy minimized applying MMFF94x force field [64] and generalized Born model and keeping the coordinates of protein backbone atoms fixed. Successively all the enzyme structures were superimposed with MOE.

Ligands were described with the general Amber force field GAFF [65]. RESP charges for compound **1** were calculated, while the parameters of Ulf Ryde were taken for NADPH (charge -4) (<http://www.teokem.lu.se/~ulf/>).

Molecular docking

Molecular docking calculations were performed for all inhibitors of Table 2. GOLD program [58, 59] follows a genetic algorithm (GA) search strategy which enables flexible docking of small molecules, so no conformational

search was employed to the ligand structures before docking. All ligands were docked in 50 independent GA runs. The GA default parameters were set as suggested by the GOLD authors. The automatic active-site detection was switched on. The annealing parameters of fitness function were set at 3.5 Å for hydrogen bonding and 6.5 Å for van der Waals interactions. In order to extend the ligand fitting points for the matching of hydrophobic regions to carbons, halogens and non-polar sulfur atoms the LIGAND_FITPTS_SELECTION EXTENDED_HAL_S function of the gold_parameters.dat was switched on. The GOLDScore [58] was used as scoring function to rank the docking poses.

Molecular dynamics simulations

Classical molecular dynamics simulations and binding free energy estimations were carried out using the Amber 9.0 suite [66] and the AMBER99SB force field [67]. The simulation system was set up using the xLeap program of the AMBER suite. The simulation systems were surrounded by a truncated octahedral box of TIP3P [68] water molecules of 9 Å radius from the protein. Na⁺ counter ions were added to neutralize the system. Prior to the free MD simulations, the simulation systems were energy minimized for 1,000 steps of steepest descent followed by 10,000 steps of conjugate gradient optimization. The equilibration process was carried out with the program PMEMD on the NVT ensemble using the following procedure [69–71]: the simulation system was heated during 200 ps from 0 to 200 K at constant volume conditions. This temperature was held for additional 100 ps and afterwards raised to 300 K during 200 ps at constant volume. Protein heavy atoms, NADPH, and compound **1** were kept constraint, and the force constant was gradually reduced from 10 to 1 kcal/mol/Å². The final coordinates of the temperature equilibration routine were relaxed without restraint for 100 ps and used for the MD production run, performed at NPT physical conditions and without restraints. The total simulation length differed for the various complexes ranging from 10 to 14 ns. Temperature regulation was done at 1.0 atm of pressure by using a Langevin thermostat with a coupling constant of $\tau_{\text{autp}} = 1.5$ ps [72, 73]. The time step of the free MD simulations was 2 fs, with a cutoff of 10 Å for the non-bonded interaction, and SHAKE [74] was employed to keep all bonds involving hydrogen atoms rigid. Electrostatic interactions were computed using the Particle Mesh Ewald method [69]. All simulations were carried out in periodic boundary conditions.

The analysis of the trajectories of the MD simulations was performed with the PTRAJ module of AMBER, the MMTSB tool set (<http://mmtsb.org>) [75] and using visual

molecular dynamics (VMD) (<http://www.ks.uiuc.edu/Research/vmd/>) [76]. Hydrogen bonds, expressed in terms of % occupancy for different time intervals, were measured with the VMD H-Bonds plugin version 1.1, setting the donor–acceptor distance to 3.5 Å and using as angle cutoff 25°. Salt bridges were monitored with the VMD Salt Bridges plugin version 1.0 using as oxygen–nitrogen cut-off 3.5 Å. The PLUMED collective variable analysis tool version 1.0 integrated in VMD was used to measure the motions of the residues over a trajectory with respect to C $_{\alpha}$ of Met147 (or as for D2 with respect to C $_{\alpha}$ of Leu149). Graphical presentation of the results was prepared with PyMOL (www.pymol.org) and Excel programs.

Free energy calculations using the MM-PBSA method

Conventional MM-PBSA [46, 47] and normal-mode NMODE [48] calculations were performed using the AMBER 9 suite. The electronic and van der Waals energies were calculated by the *sander* module. The polar solvation energy was calculated with the finite-difference PB equation solver by using AMBER toolset.

For each stable sector of the MD trajectories **D1–D4b** longer than 4 ns (i.e. with an all-atom RMSD variation of less than ± 0.25 Å from the mean value of the selected timeframe; for **D4b** a variation of ± 0.35 was chosen) snapshots were collected every 30th frame (every 30 ps) and used to estimate the binding free energy by means of MM-PBSA methods and NMODE. The values reported in Table 1a correspond to the following segments: 1fdtA (**D1**) 6.5–11.5 ns, 1iol (**D2**) 2.5–7 ns, 1i5r (**D3**) 6.5–11 ns, 1fdtB (**D4a**) 6–12 ns, and 1fdtB (**D4b**) 7–12 ns. The normal mode analysis was performed to estimate the vibrational component of the entropy. Conjugate gradient and then Newton–Raphson minimizations until the root mean square of the elements of gradient vector was less than 5×10^{-4} kcal/mol were carried out in absence of solvent. A distance-dependent dielectric constant was used to mimic solvent screening. Frequencies of the vibrational modes were computed at 300 K for these minimized structures and using a harmonic approximation of the energies. Due to the high computational demand, only snapshots taken every 100th frame of each trajectory were used to estimate -TAS.

MEP calculation

For compounds **1–5** geometry optimization was performed employing the B3LYP hybrid functional [56] in combination with a 6–311++G** basis set using the package Gaussian03 [77]. The molecular electrostatic potential (MEP) maps were plotted using GaussView, version 3.0 [78], the 3D molecular graphics package of Gaussian.

These electrostatic potential surfaces were generated by mapping the 6–311++G** electrostatic potentials onto the molecular electron density (isovalue = 0.004 electron/Å) and the ESP values on the surface were color-coded ranging from –0.01 (red) to +0.04 (blue) Hartree.

Acknowledgments MN is grateful for his co-tutelle PhD between the University of Bologna and Saarland University of which this work was a substantial part.

References

1. Calkin AC, Sudhir K, Honisett S, Williams MR, Dawood T, Komesaroff PA (2002) J Clin Endocrinol Metab 87:5072–5075
2. Grodstein F, Manson JE, Stampfer MJ (2001) Ann Intern Med 135:1–8
3. Lew R, Komesaroff P, Williams M, Dawood T, Sudhir K (2003) Circ Res 93:1127–1133
4. Travis RC, Key TJ (2003) Breast Cancer Res 5:239–247
5. Dizerega GS, Barber DL, Hodgen GD (1980) Fertil Steril 33:649–653
6. Sasano H, Miki Y, Nagasaki S, Suzuki T (2009) Pathol Int 59:777–789
7. Labrie F, Luu-The V, Lin SX, Labrie C, Simard J, Breton R, Bélanger A (1997) Steroids 62:148–158
8. Labrie F, Luu-The V, Lin SX, Simard J, Labrie C, El-Alfy M, Pelletier G, Bélanger A (2000) J Mol Endocrinol 25:1–16
9. Prehn C, Möller G, Adamski J (2009) J Steroid Biochem Mol Biol 114(1–2):72–77
10. Persson B, Kallberg Y, Bray JE, Bruford E, Dellaporta SL, Favia AD, Duarte RG, Jörnvall H, Kavanagh KL, Kedishvili N, Kisiela M, Maser E, Mindnich R, Orchard S, Penning TM, Thornton JM, Adamski J, Oppermann U (2009) Chem Biol Interact 178:94–98
11. Husen B, Huhtinen K, Poutanen M, Kangas L, Messinger J, Thole H (2006) Mol Cell Endocrinol 248:109–113
12. Lamminen T, Saloniemi T, Huhtinen K, Koskimies P, Messinger J, Husen B, Thole H, Poutanen H (2009) Mol Cell Endocrinol 301:158–162
13. Gunnarsson C, Hellqvist E, Stal O (2005) Br J Cancer 92:547–552
14. Gunnarsson C, Olsson BM, Stal O (2001) Cancer Res 61:8448–8451
15. Miyoshi Y, Ando A, Shiba E, Taguchi T, Tamaki Y, Noguchi S (2001) Int J Cancer 94:685–689
16. Suzuki T, Moriya T, Ariga N, Kaneko C, Kanazawa M, Sasano H (2000) Br J Cancer 82:518–523
17. Delvoux B, Groothuis P, D’Hooghe T, Kyama C, Dunselman G, Romano A (2009) J Clin Endocrinol Metab 94:876–883
18. Jansson A (2009) J Steroid Biochem Mol Biol 114:64–67
19. Saloniemi T, Järvensivu P, Koskimies P, Jokela H, Lamminen T, Ghaem-Maghami S, Dina R, Damdimopoulou P, Mäkelä S, Perheentupa A, Kujari H, Brosens J, Poutanen M (2010) Am J Pathol 176:1443–1451
20. Blomquist CH, Bonenfant M, McGinley DM, Posalaky Z, Lakatua DJ, Tuli-Puri S, Bealka DG, Tremblay Y (2002) J Steroid Biochem Mol Biol 81:343–351
21. Kasai T, Shozu M, Murakami K, Segawa T, Shinohara K, Nomura K, Inoue M (2004) J Clin Endocrinol Metab 89:5661–5668
22. Nagasaki S, Miki Y, Akahira J, Suzuki T, Sasano H (2009) Ann N Y Acad Sci 1155:25–32

23. Baston E, Paluszczak A, Hartmann RW (2000) *Eur J Med Chem* 35:931–940
24. Miller WR, Bartlett JM, Canney P, Verrill M (2007) *Breast Cancer Res Treat* 103:149–160
25. Bush NJ (2007) *Semin Oncol Nurs* 23:46–54
26. Adamo V, Iorfida M, Montalto E, Festa V, Garipoli C, Scimone A, Zanghi M, Caristi N (2007) *Ann Oncol* 18(Suppl 6):vi53–vi57
27. Labrie F (1991) *Mol Cell Endocrinol* 78:C113–C118
28. Day JM, Foster PA, Tutill HJ, Parsons MF, Newman SP, Chander SK, Allan GM, Lawrence HR, Vicker N, Potter BV, Reed MJ, Purohit A (2008) *Int J Cancer* 122:1931–1940
29. Brožič P, Lanišnik Rižner T, Gobec S (2008) *Curr Med Chem* 15:137–150 (and references therein cited)
30. Poirier D (2009) *Anticancer Agents Med Chem* 9:642–660
31. Day JM, Tutill HJ, Purohit A, Reed MJ (2008) *Endocr Relat Cancer* 15:665–692
32. Messinger J, Hirvela L, Husen B, Kangas L, Koskimies P, Pentikainen O, Saarenketo P, Thole H (2006) *Mol Cell Endocrinol* 248:192–198
33. Lilienkampf A, Karkola S, Alho-Richmond S, Koskimies P, Johansson N, Huhtinen K, Vihko K, Wähälä K (2009) *J Med Chem* 52:6660–6671
34. Allan GM, Vicker N, Lawrence HR, Tutill HJ, Day JM, Huchet M, Ferrandis E, Reed MJ, Purohit A, Potter BV (2008) *Bioorg Med Chem* 16:4438–4456
35. Frotscher M, Ziegler E, Marchais-Oberwinkler S, Kruchten P, Neugebauer A, Fetzler L, Scherer C, Müller-Vieira U, Messinger J, Thole H, Hartmann RW (2008) *J Med Chem* 51:2158–2169
36. Marchais-Oberwinkler S, Kruchten P, Frotscher M, Ziegler E, Neugebauer A, Bhoga U, Bey E, Müller-Vieira U, Messinger J, Thole H, Hartmann RW (2008) *J Med Chem* 51:4685–4698
37. Marchais-Oberwinkler S, Frotscher M, Ziegler E, Werth R, Kruchten P, Messinger J, Thole H, Hartmann RW (2009) *Mol Cell Endocrinol* 301:205–211
38. Oster A, Klein T, Werth R, Kruchten P, Bey E, Negri M, Marchais-Oberwinkler S, Frotscher M, Hartmann RW (2010) *Bioorg Med Chem* 18:3494–3505
39. Oster A, Hinsberger S, Werth R, Marchais-Oberwinkler S, Frotscher M, Hartmann RW (2010) *J Med Chem* 53:8176–8186
40. Bey E, Marchais-Oberwinkler S, Kruchten P, Frotscher M, Werth R, Oster A, Algul O, Neugebauer A, Hartmann RW (2008) *Bioorg Med Chem* 16:6423–6435
41. Bey E, Marchais-Oberwinkler S, Werth R, Negri M, Al-Soud YA, Kruchten P, Oster A, Frotscher M, Birk B, Hartmann RW (2008) *J Med Chem* 51:6725–6739
42. Bey E, Marchais-Oberwinkler S, Negri M, Kruchten P, Oster A, Werth R, Frotscher M, Birk B, Hartmann RW (2009) *J Med Chem* 52:6724–6743
43. Al-Soud YA, Bey E, Oster A, Marchais-Oberwinkler S, Werth R, Kruchten P, Frotscher M, Hartmann RW (2009) *Mol Cell Endocrinol* 301:212–215
44. Negri M, Recanatini M, Hartmann RW (2010) *PLoS ONE* 5:e12026
45. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) *Nucl Acids Res* 28:235–242
46. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatham TE III (2000) *Acc Chem Res* 33:889–897
47. Srinivasan J, Cheatham TE III, Cieplak P, Kollman PA, Case DA (1998) *J Am Chem Soc* 120:9401–9409
48. Case DA (1999) In: Thorpe MF, Duxbury PM (eds) *Rigidity theory and applications*. Plenum, New York, pp 329–344
49. Hou T, Wang J, Li Y, Wang W (2011) *J Chem Inf Model* 51:69–82
50. Rastelli G, Del Rio A, Degliesposti G, Sgobba M (2010) *J Comput Chem* 31:797–810
51. Muzzioli E, Del Rio A, Rastelli G (2011) *Chem Biol Drug Des* doi:10.1111/j.1747-0285.2011.01140.x
52. Jin JZ, Lin SX (1999) *Biochem Biophys Res Commun* 259:489–493
53. Mobley DL, Chodera JD, Dill KA (2007) *J Chem Theory Comput* 3:1231–1235
54. Hohenberg P, Kohn W (1964) *Phys Rev* 136:B864–B871 39
55. Kohn W, Sham LJ (1965) *Phys Rev* 140:A1133–A1138
56. Becke AD (1993) *J Chem Phys* 98:5648–5652
57. Gasteiger J, Li X, Rudolph C, Sadowsky J, Zupan J (1994) *J Am Chem Soc* 116:4608–4620
58. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) *J Mol Biol* 267:727–748
59. Verdonk ML, Cole JC, Hartshorn MJ, Murray CW, Taylor RD (2003) *Proteins* 52:609–623
60. Mauldin RV, Carroll MJ, Lee AL (2009) *Structure* 17:386–394
61. Molecular Operating Environment MOE 2008.10; C.C.G.I.M.: Quebec, Canada, 2008; <http://www.chemcomp.com>
62. Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC (2010) *Acta Cryst D* 66:12–21
63. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB 3rd, Snoeyink J, Richardson JS, Richardson DC (2007) *Nucl Acids Res* 35:W375–W383
64. Halgren TA (1999) *J Comput Chem* 20:730–748
65. Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) *J Comput Chem* 25:1157–1174
66. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Pearlman DA, Crowley M, Walker RC, Zhang W, Wang B, Hayik S, Roitberg A, Seabra G, Wong KF, Paesani F, Wu X, Brozell S, Tsui V, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Beroza P, Mathews DH, Schafmeister C, Ross WS, Kollman PA (2006) *AMBER 9*. University of California, San Francisco
67. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) *Proteins: Struct Funct Bioinf* 65:712–725
68. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) *J Chem Phys* 79:926–935
69. Darden T, Perera L, Li L, Pedersen L (1999) *Structure* 7:R55–R60
70. Essman U, Perela L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) *J Chem Phys* 103:8577–8592
71. Sagui C, Darden T (1999) *Annu Rev Biophys Biomol Struct* 28:155–179
72. Adelman SA, Doll JD (1976) *J Chem Phys* 64:2375–2388
73. Feller SE, Zhang Y, Pastor RW, Brooks BR (1995) *J Chem Phys* 103:4613–4621
74. Ryckaert JP, Ciccotti G, Berendsen JC (1977) *J Comput Phys* 23:327–341
75. Feig M, Karanicolas J, Brooks CL 3rd (2004) *J Mol Graph Model* 22:377–395
76. Humphrey W, Dalke A, Schulten K (1996) *J Mol Graph* 14: 33–38
77. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O,

Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW,

Gonzalez C, Pople JA (2004) Gaussian 03, revision C.02. Gaussian Inc., Wallingford

78. Dennington I, Roy KT, Millam J, Eppinnett K, Howell WL, Gilliland R (2003) GaussView, version 3.0; Semichem Inc., Shawnee Mission, KS