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

Matthias Negri,[†], Maurizio Recanatini,[‡] and Rolf W. Hartmann [†]*

[†] *Pharmaceutical and Medicinal Chemistry, Saarland University & Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Campus C2.3, D-66123 Saarbrücken, Germany.*

[‡] *Department of Pharmaceutical Sciences, University of Bologna, Via Belmeloro, 6, I-40126 Bologna, Italy*

* To whom correspondence should be addressed: Prof. Dr. Rolf W. Hartmann. Pharmaceutical and Medicinal Chemistry, Saarland University, Campus C2.3, D-66123 Saarbrücken, Germany; Phone: +49-681-30270300. Fax: +49-681-30270308. E-mail: rwh@mx.uni-saarland.de. Homepage: http://www.PharmMedChem.de

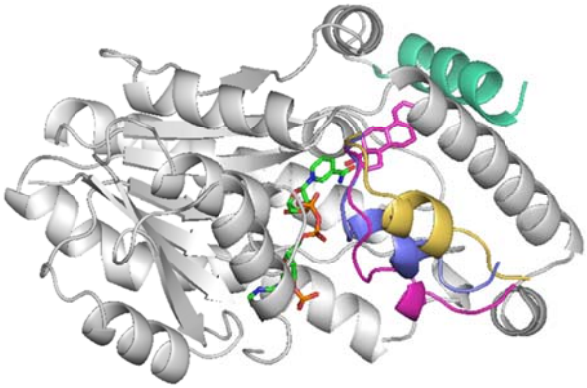

Fig ESM1 Rigid (grey) and flexible regions (coloured) of 17β-HSD1 resulting from the analysis of the crystal structures. The three βFαG'-loop conformations resulting from backbone clustering are coloured in magenta (cl3), purple (cl2) and yellow (cl1) (Negri M, Recanatini M, Hartmann RW (2010) PLoS ONE 5:e12026)

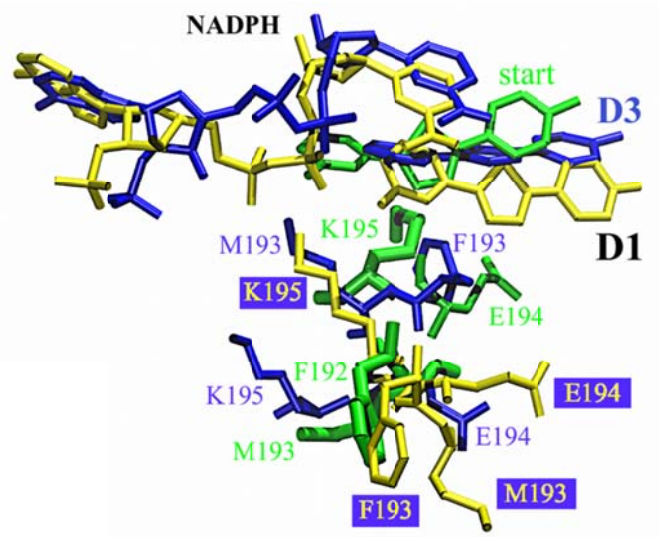

Fig ESM2 Schematic representation of the βFαG'-loop residues, compound **1** and NADPH in the starting conformation of **D1** (green), and of the final complex of **D1** (yellow) and of **D3** (blue)

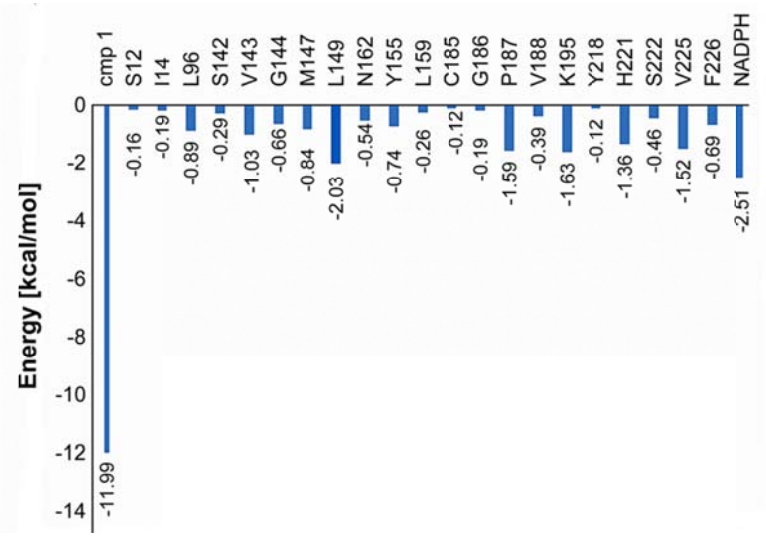

Fig ESM3 Energy decomposition for the MD **D1**. The relative energy contributions of the amino acids, NADPH and compound **1** to the estimated enthalpic term (via molecular mechanics General Born surface area (MM-GBSA) methods) are shown, supportive the positive influence of compound **1** binding (largest enthalpy contribution).

	residue	0-1 ns	1-3 ns	3-6 ns	6-9 ns	9-12 ns	average (0-12 ns)
alternative binding mode							
D1	Gly144	3%	0	0	0	0	0
	Met147	4.5%	0.5%	0	0	0	0.5%
	Lys193	0	4%	4%	5%	6%	4%
	His221	0	59%	78%	79%	78%	69%
	NADPH	94%	84%	83%	86%	80%	84%
D2	Gly144	0	0	0	1%	1%	0.5%
	Asn152	0	1.5%	8%	17%	16%	11%
	NADPH	68%	95%	45%	33%	87%	62%
D3	Met147	0	0	1%	0	0	0.5%
	Met193	0	0	0	1%	0	0.5%
	NADPH	95%	95%	97%	89%	93%	94%
steroidal binding mode							
D4a	Ser142	25%	17%	0	0	0	5%
	Asn152	0	12%	15%	36%	80%	43%
	Tyr155	3%	0%	0	0	0	0
	His221	10%	0	0	0	0	1%
	Arg258	0	1%	12.5%	2%	1%	3%
D4b	Ser142	0	5%	2.5%	0	0	1.5%
	Asn152	4.5%	0.5%	0	0	0	0.5%
	Tyr155	9%	3%	1%	0	0	2%
	Tyr218	1%	2.5%	2%	0	0	1%
	His221	24%	15%	10%	21%	16%	19%
	Tyr275	5%	1%	0	0	0	1%
	Glu282	0	0	0	2%	1%	1%

Fig ESM4 Hydrogen bonds of compound 1 expressed in terms of % occupancy and measured for different trajectory segments of the five MD simulations **D1**-**D4b**.

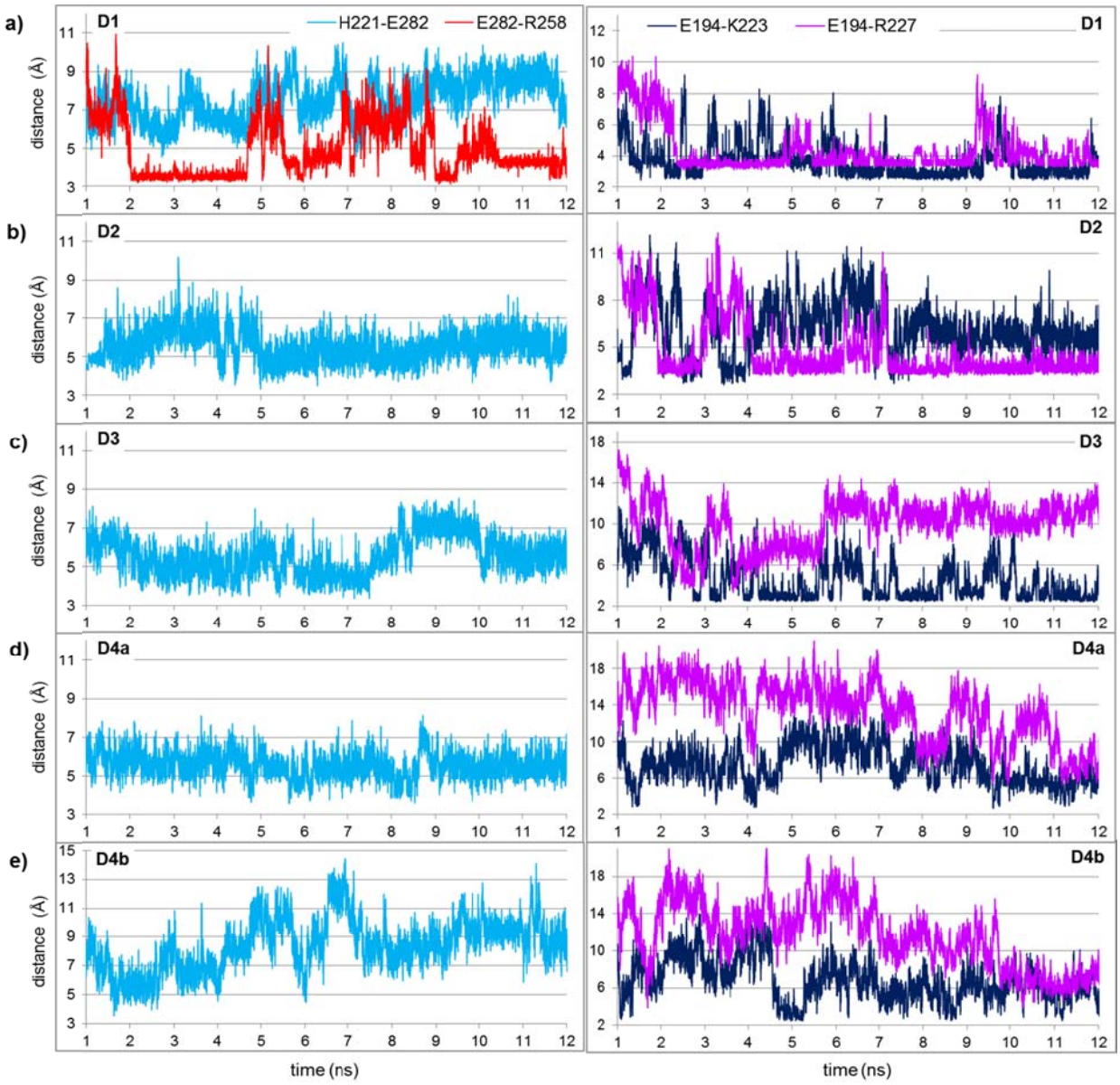

Fig ESM5 Important salt-bridges monitored for **D1**-**D4b (a-e)**.The plots on the left show the distance (Å) between **H221/E282** (marine) and **R259/E282** (red; present only in **D1**) and regarding the "open/closed" state of gate 3 we observed that: 1) open gate when no H221/E282 and R258-E282; 2) closed gate 3 when R258-E282.
The plots on the right show the distance between **E194/K223** (dark blue) and **E194/R227** (magenta) – considerations to the "open/closed" state of gate 2 are: 1) closed gate 2 when **E194/K223** and **E194/R227**; 2) open gate 2 when no salt bridge or only **E194/K223** (loop is then centrally stabilized).

Distance collective variables (CV).

For the MD simulations **D1-D4b** we traced the distance collective variables (CV), obtained with the PLUMED plugin 1.0 in VMD 1.9, over the simulation timeframes These CV are a quantitative measure of the motions of key residues (i.e. F192 and F226 - gate 1 and 2, and H221, R258, and E282 - gate 3).with respect to a reference atom (or group of atoms – center of mass) and to monitor how these movements influence active site topology and ligand binding. As reference we chose the C_α of Met147 (**D1, D3, D4a, D4b**) and of Leu149 (**D2**), because these atoms showed a stable trajectory (low RMSD; see Fig 5 and 8) and because they are placed in the central lower part of the active site wall facing the flexible βFαG'-loop. This last aspect allowed us to monitor the effects on gate 1, 2 and 3, correlating their opening and closing with the increase/decrease in RMSD of a key residue, Aiming to study the aspect of ligand binding we chose as references the center of mass of all atoms of compound 1 or of its three rings singularly (reg I, II, III).

In the legends to Fig ESM6, ESM7, ESM9, ESM10 and ESM11 we also interpret the data with respect to striking motions for each MD simulation and to their impact on the active site topology.

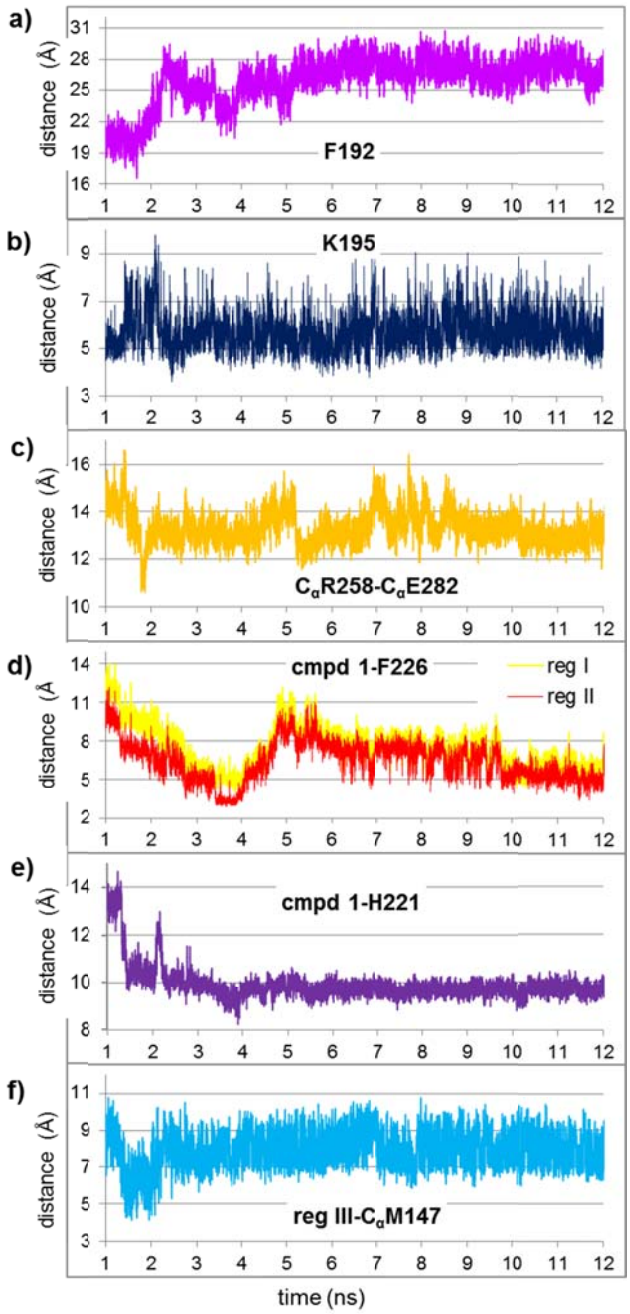

Fig ESM6 Collective variables (CV) for **D1** plotted over the simulation time.
a) C_α Met147 - center of mass of **Phe192.** The phenyl ring rotates away from the active site center (Met147) toward the adenosine of NADPH – aiming to occlude gate 1 and stabilizing the cofactor.
b) C_α Met147 - center of mass of **Lys195** (sidechain). In the last part of the MD the average increases of 1 Å (6.5 Å) - in concomitance **Lys**195 gradually shifts to the cofactor phosphates closing gate 1.
c) C_α Arg258 - C_α Glu282. After few ns the RMSD decreases in concomitance with the formation of the ion-pair R258-E282 (Fig ESM5a) - thereby these two residues occlude gate 3;
d) Center of mass of compound **1**- center of mass of **Phe226**. The initial decrease in RMSD and the stable plateau in the second half of the trajectory suggest a slow, gradual induced-fit of the protein to the inhibitor and a strong binding;
e) Center of mass of compound **1**- center of mass of **His221** (imidazole). The drop in RMSD coincides with the loss of the salt-bridge with Glu282 as well as with the hydrogen bond formation with compound 1 (see Fig ESM4);
f) C_α Met147 - **reg III.** The drop in RMSD for the p-hydroxyphenyl-ring of 1 (region III) accounts for the initial motion towards Met147, while its progressive increase reflects the shift toward αG-helix and His221;

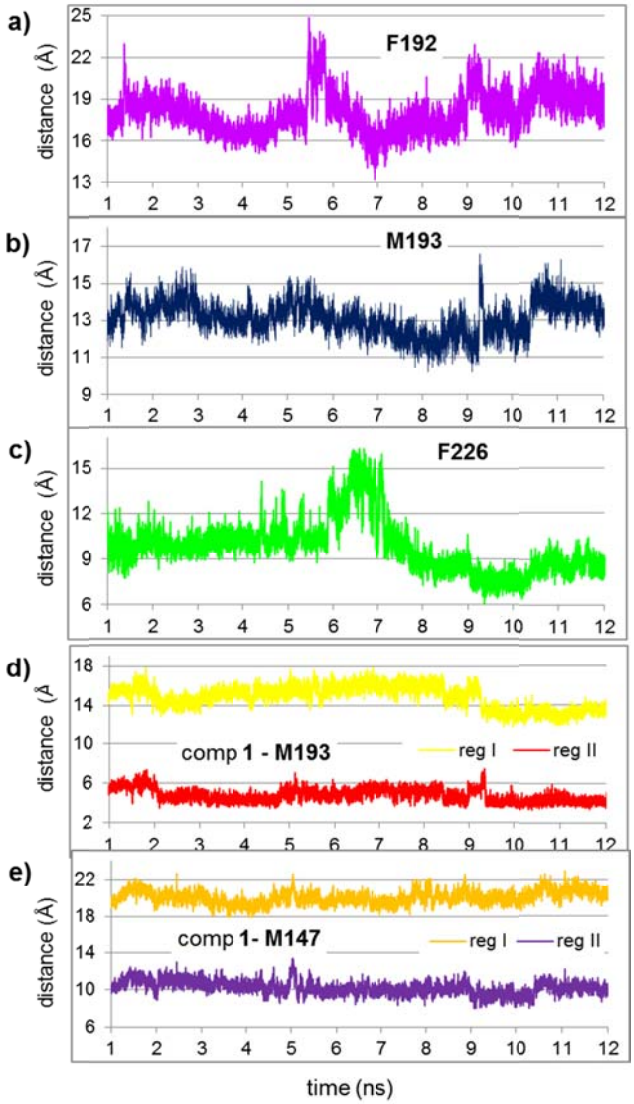

Fig ESM7 Collective variables (CV) for **D2** plotted over the simulation time.
a) C_α Leu149 - center of mass of **Phe192** (phenyl ring). The distance is much shorter then in **D1** and reflects a more occluded gate 1. The large excursus in Å at 5.5-6 ns correlates well with the RMSD increases in Fig 5b and d – F192 stabilizes NADPH and occludes gate 1;
b) C_α Leu149 - center of mass of **Met193** (sidechain). The progressive decrease in distance corresponds to a gradual shift of the loop axis toward the αG-helix and the center of the active site – thereby stabilizing the inhibitor;
c) C_α Leu149 - center of mass of **Phe226** (sidechain). After a stable plateau in the first 6 ns a high peak follows due to the loop axis motion. Then F226 rotates inside and reaches a new plateau – it closes gate 2 and stabilizes **1**;
d) C_α Met193 – **reg I** and **II**. After an initial drop the distance between M193 and compound **1** remains stable for the first 9 ns, reflecting the tight interaction M193 with the central ring of **1** (**reg II**; red). In the last 3 ns M193 moves away from L149, as shown in **b**), towards **reg I** – shielding the cofactor binding site and closing gate 1);
e) C_α Met147 – **reg I** and **II**. The distance profile of **reg I** and **II** (compound **1**) to M147 is constant, similar to the one for M193, which forms the opposite wall of the active site, see **d**) – 1 is sandwiched between V143, M147, L149, and F259 (buried wall) and F192, M193, F226 (FG-segment; "solvent exposed" wall). See also Fig ESM8.

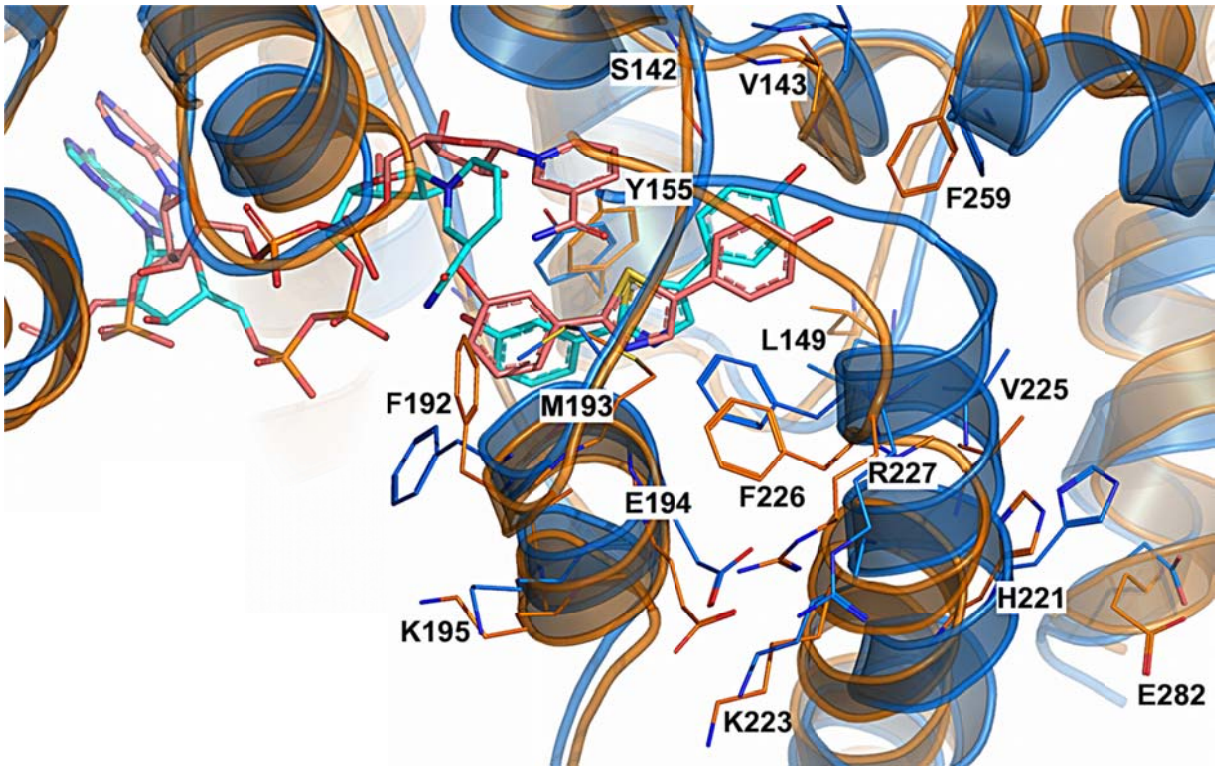

Fig ESM8 Analysis of the MD simulations **D2**. This overlay of a snapshot from the first stable plateau (2-7 ns; orange) and of the final structure (10-12 ns; marine blue) shows the inwards shift of the whole FG-segment (βFαG-loop and αG'-helix) narrowing the substrate binding site. F226 shifts toward the inhibitor (reg II; central thiazole), while F192 moves toward the cofactor, blocking gate 1. Notably, the αG'-helix and the C-terminal helix moved in a concerted way with the salt-bridge between H221 and E282 maintained.

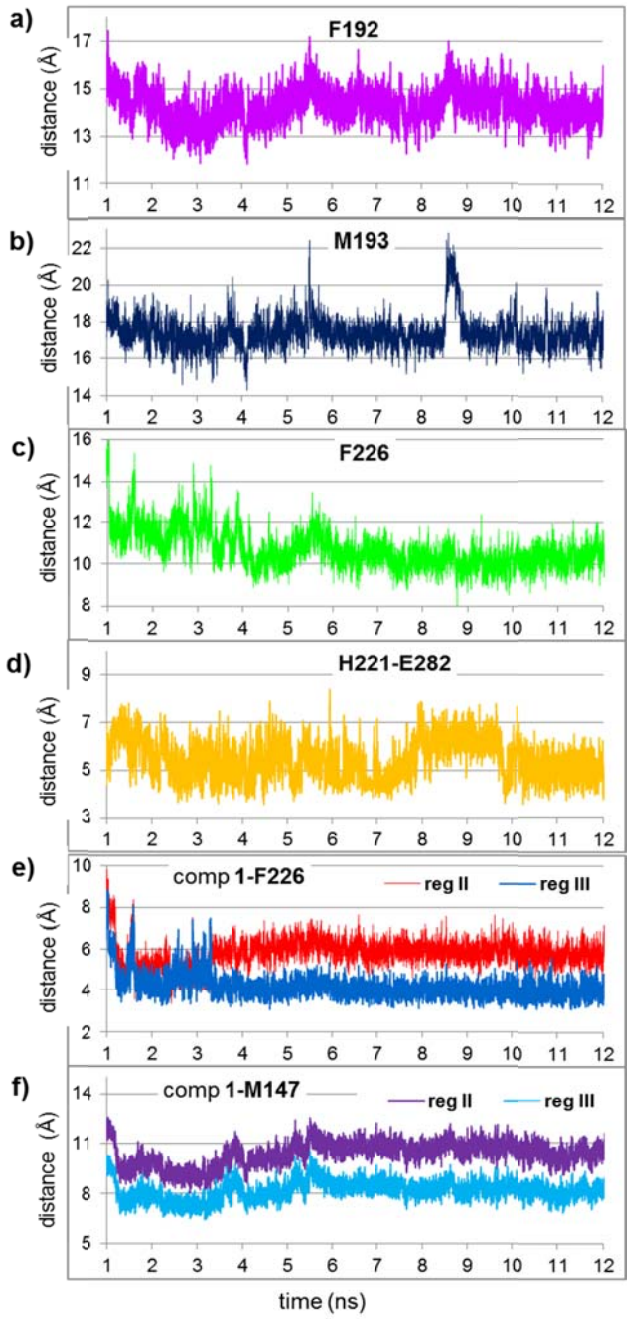

Fig ESM9 Collective variables (CV) for **D3** plotted over the simulation time.
a) C_α Met147 - center of mass of **Phe192** (phenyl ring). The distance to M147 initially drops from 17 to 15 Å, but then it remains stable – inward motion of F192 to stabilize compound **1**.
b) C_α Met147 - center of mass of **Met193** (sidechain). As for F192 the distance to the center of the active site initially decreases (from 20 to 17 Å), but still it is far away and appears to be related to the loop formation. Then it remains stable except for a few peaks, which are also associated to the βFαG'-loop motions (Fig 8c)
c) C_α Met147 - center of mass of **Phe226** (phenyl ring).F226 moves into the active site in the first 6 ns, then it reaches a stable position where it interacts via π-π stacking with **reg III** of compound **1**, in particular **e**).
d) ion-pair **His221**/**Glu282.** The salt-bridge is relatively stable suggesting a conserved topology for the C-terminal part.
e) Center of mass of **Phe226** - **reg II** and **III**. In the first ns F226 comes closer to reg II and III, stabilizing the inhibitor via π-π stacking see c).
f) C_α Met147 - **reg II** and **III.** After the induce-fit at the beginning, which brings 1 closer to M147, reg II and III shift again toward the FG-segment, contributing to stabilize the inhibitor involved in π-stacking with D192 and F226.

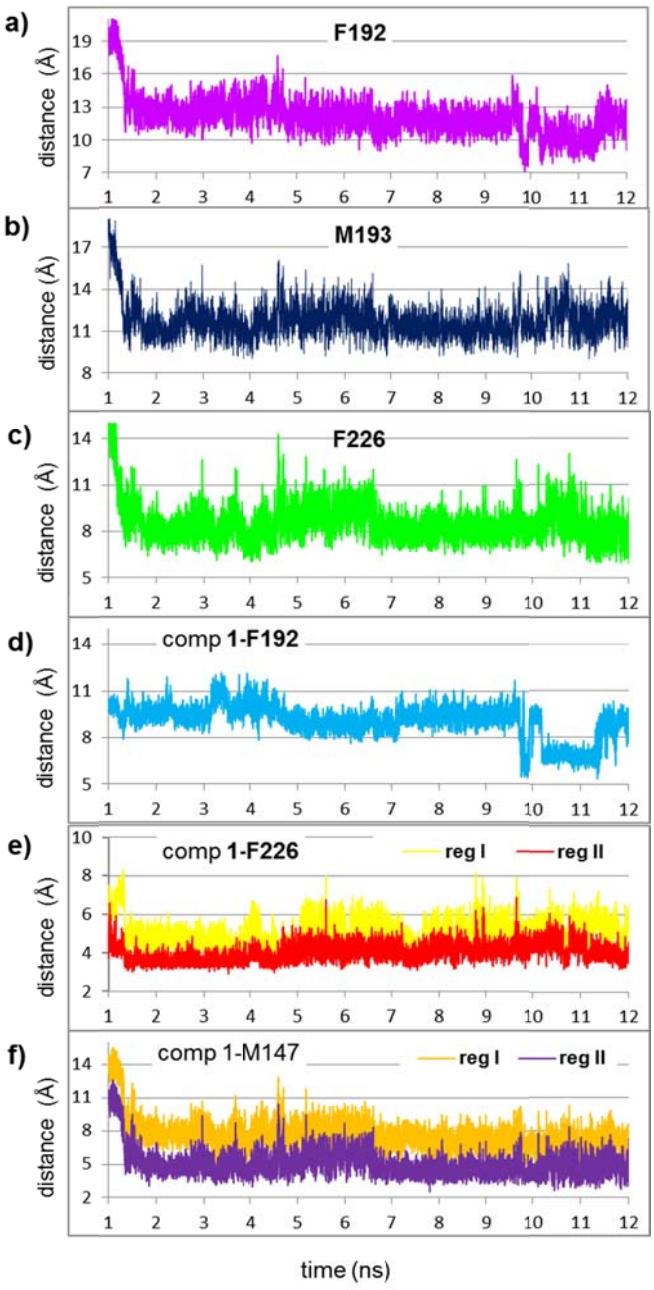

Fig ESM10 Collective variables (CV) for **D4a** plotted over the simulation time.
a) C_α Met147 - center of mass of **Phe192** (phenyl ring). The distance to M147 decreases in the first ns from 19 to 13 Å, where it remains for several ns before protruding further into the active site. Thereby **1** interacts with F192, which contributes to hold the hydrogen bond between **reg I** and Asn152;
b) C_α Met147 - center of mass of **Met193** (sidechain). This distance profile as well as that of F192 presents an initial drop followed by a long lasting stable plateau with final fluctuations. The drop can be associated with the movement of the loop into the active site as shown by the decrease in distances of F192 and M193 with respect to M147, centrally placed in the active site;
c) C_α Met147 - center of mass of **Phe226** (phenyl ring). F226 follows the same scheme as F192 and M193. The shift of the whole FG-segment toward the active site center leads to a narrow active site and to induced-fit motions of the enzyme toward the inhibitor.
d) Center of mass of **1**- center of mass of **Phe192**. In the final two ns compound **1** and F192 are close enough to interact. This decrease in distance is concomitant with the increase in loop RMSD in Fig 8c;
e) Center of mass of **Phe226** – **reg I** and **II**. Both the *m*-OH-phenyl-ring and the central heterocycle X show an overall stability. In the last two ns F226 drops on X stabilizing it via π-stacking.
f) C_α Met147 - **reg I** and **II**. Same as for e). **1** is pulled toward the active site wall opposite to the loop, where it is involved in a hydrophobic network with V143, M147, L149, F259, etc. Thereby, van der Waals interactions with M147 contribute to a stable inhibitor binding.

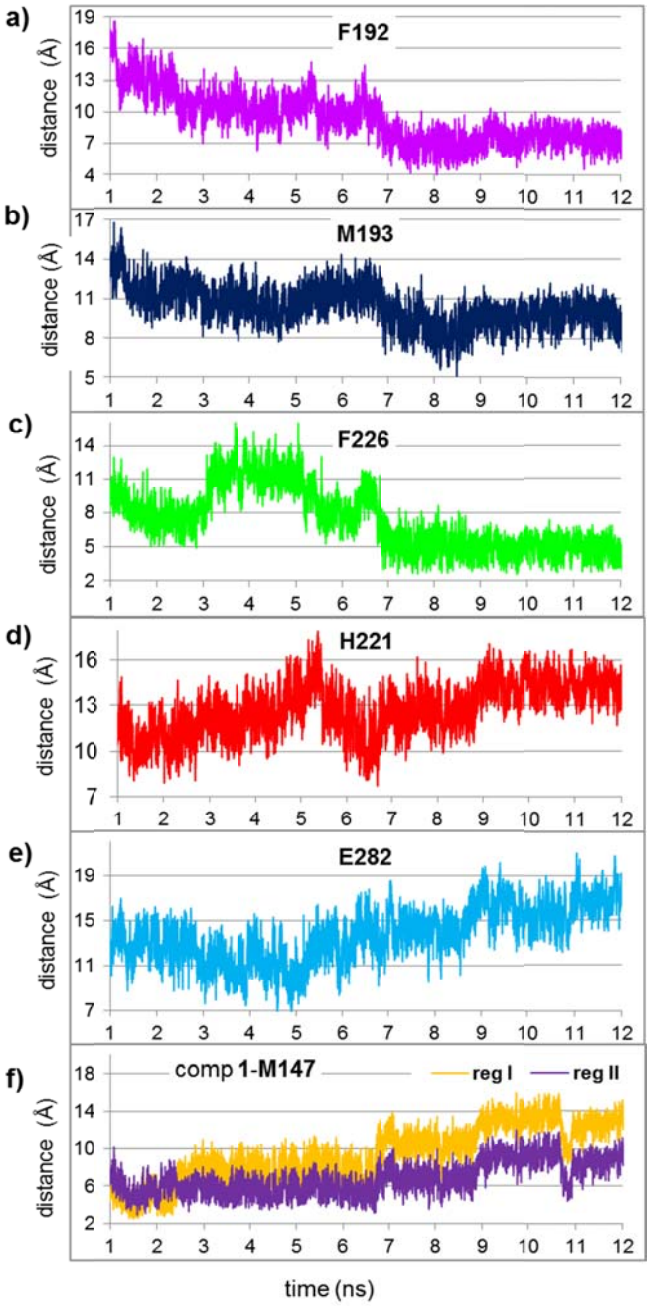

Fig ESM11 Collective variables (CV) for **D4b** plotted over the simulation time.
a) C_α Met147 - center of mass of **Phe192** (phenyl ring). The first half of the trajectory is characterized by the shift of F192 toward the active site center shown by the drop in distance of about 13 Å. After 8 ns F192 finds a stable average position, which it maintains until the end of the trajectory. At this stage the loop residues separate the active site in two distinct cofactor and substrate binding sites.
b) C_α Met147 - center of mass of **Met193** (sidechain). A gradual decrease of the distance of about 6 Å occurs in the first half of the trajectory. This, again, shows the protrusion of the loop in the active site.
c) C_α Met147 - center of mass of **Phe226**. Large conformational rearrangements bring F226 deep into the active site. After them the compound **1** changes its orientation (7ns), with a consequent stabilization of the pose of M147, F192, M193, and F226 as seen in f) and Fig 8d.
d) C_α Met147 - center of mass of **His221**. As shown in the plot H221 moves away from the center (M147) in the first nine ns, ending solvent exposed and edged out of gate 3. The short plateau at the end coincides with one existing for E282 e)
e) C_α Met147 - center of mass of **Glu282**. Also E282 turns away from the center and edges outward.
f) C_α Met147 - **reg I** and **II**. Notably the distance of **reg I** and **II** from M147 rises consistently over the timeframes. This accounts for the "expulsion" of **1** observed in this MD, with the *m*-OH-phenyl involved in a hydrogen bond with H221 and the *p*-OH-phenyl ring solvent exposed at gate 3.