PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Activation helix orientation of estrogen receptor is mediated by receptor dimerization: evidence from molecular dynamics simulations

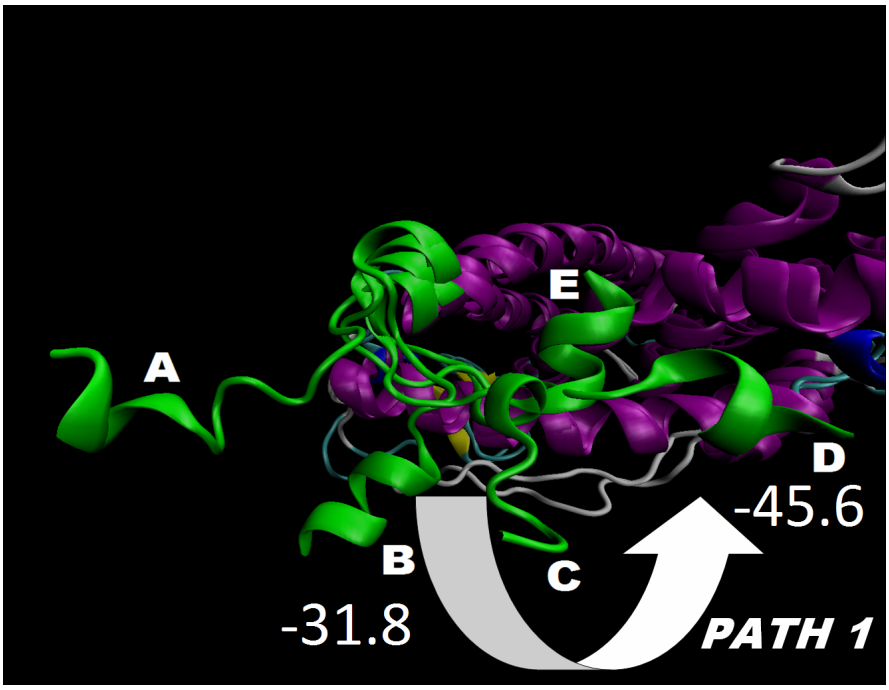
Filip Fratev* [1, 2]

[1] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, Acad. G. Bonchev Str. Block 105, 1113 Sofia, Bulgaria

[2] Micar21 Ltd., Persenk Str. 34B, 1407 Sofia, Bulgaria

[*] *Corresponding author*: e-mail: fratev@biomed.bas.bg; Tel: +359876309111;

Keywords: Estrogen receptor (ERα) · helix 12 · Accelerated Molecular Dynamics (aMD) · LBD ·Nuclear receptors (NR)



ABSTRACT:

In the recent years the nuclear receptors (NR) dynamics have been extensively studied by various approaches. However, the transition path of helix 12 (H12) to an agonist or an antagonist conformation and the exchange pathway between these states is not clear yet.

A number of accelerated molecular dynamics (aMD) runs were performed on both an ERα monomer and a homodimer with a total length of 2.2 μs. We have been able to sample reasonably well the H12 conformational landscape, to reproduce precisely both the agonist and the antagonist conformations, starting from an unfolded position, and to describe the transition path between them even in the presence of an agonist ligand. These conformations were the most prevalent, suggesting that the extended H12 state is not likely to exist and also that the natural ERα H12 position might exist in both the agonist or antagonist states. Remarkably, the H12 transition occurs and is regulated only in a dimer form and the proper agonist or antagonist H12 conformation can be achieved solely in one of the dimer subunits. These results clearly demonstrate that clusters of the two well known H12 states exist by themselves in the protein free energy landscape, i.e. they are not constituted directly by the ligands, and dimerization favors the switch between them. Conversely, in a monomer no transitions have been observed. Thus, the dimer formation helps the constitution of populations of discrete H12 conformational states and reshapes the conformational landscape. Further analyses have shown that these observations can be explained by specific interface and long range protein-protein interactions resulting in conformational fluctuations in helices 5 and 11. Based on these results, a new ERα activation/deactivation mechanism and, thereupon, the sequence of the binding events during receptor activity modulation have been suggested according to which the ligands control H12 conformation via alteration the inter-dimer interactions. These findings agree with the HDX and fluorescence experiments and provide an explanation on a structural basis of these data which demonstrated that the dynamics of H12 are not greatly altered upon ligand binding and large fluctuations at the end of H11 are present.

INTRODUCTION

The estrogen receptor alpha (ERα) is a member of the nuclear receptor (NR) superfamily [1]. These proteins are ligand-activated transcription factors involved in a number of biological processes, such as homeostasis, lipid metabolism, embryonic development, and cell death [2-5]. ERα consists of three functional domains. The coactivators bind to the C-terminal

ligand binding domain (LBD), where the activation function-2 (AF-2) is positioned [6]. Upon activation, the major conformational change takes place in the LBD where helix 12 (named H12 here) is repositioned to either cap the ligand binding site, as for the agonist protein structures, or to reside in a significantly different conformation, as seen in the antagonist protein structure [7].

Most of the information about NR structure-function relationships has been derived from hundreds of X-ray crystallography studies, which have captured static conformational "snapshots" of many ligand-receptor complexes. The comparison of the LBD of NRs has shown that at least three different structural conformations are present. All these conformations share a certain similarity in the binding site region, but a major difference in the H12 position [8]. The structure used to model the apo conformation of ERα (pdb id: 1a52) reveals that H12 is very flexible and when comparing this conformation to known apo structures of other NRs, it becomes obvious that there is a great resemblance [9-12]. Initially, it was commonly believed that it served as the best available model of an apo conformation of ERα LBD. However, most recent studies have indicated that this structure is presumably an artifact from the X-ray technique used [13-14].

The proper H12 conformation in an agonist form allows coactivator binding and is one of the critical steps for ERα activation, which were greatly explored in drug design [4]. It has been reported as well that the transcriptional activity correlates with the H12 position [15]. Prior to NMR experiments in the nuclear receptor field, it was thought that LBD existed only in discrete conformational states, depending on the specific ligand bound to the receptor, and the ligand-binding event shifted the conformation from one state to another (mouse-trap model) [15-20]. The latter suggestion, however, cannot explain many of the observed NR features, including the partial agonism. These also include the non classical helix 12 independent activation, such as that of the selective estrogen receptor modulators (SERMs). Furthermore, it has been proved that the H12 position is not trivially determined by the volume of the ligand and small-size ligands can be antagonists too [21-23]. Based solely on the information obtained by X-ray structures, several mechanisms have been suggested for partial agonists, including an intermediate and quasi-antagonist conformation, where coregulator interaction can switch H12 between different positions [24-26].

The NMR, HDX and fluorescence experiments have greatly enhanced the knowledge about NR's structural dynamics and partial agonism but it should be noted that they were mostly performed on other NRs and to a less extent on ERα. These data for various NRs support a different model than the one initially suggested (i.e. mouse-trap model) and was based solely

on X-ray structures, by which the apo LBD samples an ensemble of multiple conformations, and ligand binding stabilizes a subset of these conformations [27-30]. Some studies have also revealed that NR modulators can sample multiple binding modes manifesting in multiple receptor conformations in slow conformational exchange [31]. The NMR of ERβ LBD demonstrated an altered conformation in the presence of 17β-estradiol too [32]. The only available, to the best of our knowledge, NMR study for ERα has shown an overall structure stabilization upon ligand binding but has not concretisized this process to the individual LBD structural elements and in particular H12 [33]. An intrinsic tryptophan fluorescence of the ERα was used and significant fluctuations of the last part of helix 11 were identified upon ligand binding [34]. However, as it was noted before, these results cannot be directly translated to the helix 12 motion [35]. In addition, the performed HDX studies have demonstrated that the receptor dynamic profiles of some ERα-ligand complexes correlate with the pharmacological phenotype of the compound in a manner that could not be inferred from static crystal structures [35]. Notably, in the same study it has been concluded that the HDX dynamics of the ERα helix 12 exhibit no significant stabilization on binding of ligands. Recently, a similar result has been obtained for the H12 dynamics in RXRα showing that in a case of some agonist molecules only the coactivator is required to fully stabilize the activation helix [36].

It should be underlined that different NRs, even from the same class, have shown dissimilar dynamics. For instance, different HDX patterns were observed for ERβ LBD when compared to those for ERα LBD bound to the same chemical compounds, indicating that ERβ undergoes a different structural response to the same ligand when compared to ERα [37]. Thus, the above data should be carefully considered and not generalized to all NRs and in particular to ERα.

Nuclear receptors form strong dimers via their LBDs, which is crucial for the binding of cofactors and thus essential for their function as transcription factors [3]. Recently, the relationship between the dimer formation, ligand binding and H12 conformation has become more evident, but it is still poorly understood. It has been shown that apo ER LBDs exist as very stable dimers and that they are further stabilized by the addition of various ligands [38]. Estrogen antagonists provide greater kinetic stabilization of ER dimers than agonists. For instance, antagonists such as 4-hydroxytamoxifen and raloxifene induced the dimer formation of a wild type ERα LBD more potently than an agonist like 17β-estradiol (E_2), and F-domain truncation strongly depressed the ability of dimerization. As the F-domain is directly adjacent to H12, it has been proposed that this protein part would be influenced by the H12 position

[38-39]. Furthermore, it has been directly shown that the dislocation of H12 causes an antagonist dependent LBD homodimerization involving the F-domain of ERα, which facilitates dimerization, thus proving the relationship between homodimers formation and H12 conformation. A newly published X-ray structure and biochemical study revealed an allosteric mechanism whereby the ligand conformation and the cofactor binding site, including H12, of both unbound and bound receptors, are affected through the ERα and RAR homodimers interfaces [40] thus validating the negative cooperativity model for an established functional homodimer.

On the other hand, several previous studies also reported that the mutations in the ERα LBD changed H12 conformation to either an antagonist or an agonist form. For instance, the L543A/L544A and M547A/L548A mouse ERα mutants exhibited reduced basal transcription activity and lost the ability to respond to E_2 [39, 41-42], whereas L511R of H11 has been reported to result in an E2 non-active mutant because of disruption of homodimerization [43]. Further, the C381S, C417S and C530S triplet mutations produced an antagonist conformation despite the presence of a tightly bound estradiol in the ligand-binding cavity [44]. Some biophysical studies revealed the role of the phosphorylation of the clinically relevant tyrosine 537 in ERα concluding that this process has important effects on the dynamics of the helix 12 [45].

Taking together all these data, it seems that the H12 conformation can be easily perturbed and homodimer formation, similarly to an introduced mutation, might influence the H12 position and transition between different discrete states. This can explain the antagonist conformation observed in case of a smaller than natural ligand, indicating that the ligand has no single role in the H12 antagonist formation [21-23]. The latter has been recently supported by a study which has identified different NRs structural groups that might contribute to the antagonist form stabilization [46]. Thus it seems that most of the experimental data support the so called dynamic model where H12 occupies two more or less favorable states [44], and the equilibrium between the ERα conformations with the two positions of H12 (agonist or antagonist) can be perturbed by point mutations or by the presence of various other compounds.

In general, the above reviewed studies have revealed many features of H12 behavior in ERα but have also raised many questions. For instance, there is still no explanation and a structural basis of the observed by the NMR and HDX dynamics of ERα LBD, which seems to be different from other NRs such as PPARγ and even ERβ. It is also unclear how the observed conformational changes in a dimer form are provoked and whether they affect the transition

of H12 between different states. Moreover, the sequence of the binding events during ERα activation is not well understood (i.e. H12 conformational change, ligand binding, cofactor binding and homo-dimerization accompanying LBD activation) and the transition path between an agonist and an antagonist state is still not examined. There are also no available structural and dynamics data about this process at a molecular level.

In the present study we focused on the exploration of the ERα H12 conformational space in both a monomer and a dimer form. We included along the monomer MD runs extensive simulations on the dimer form as well because this step had been skipped in most of the previously performed MDs on NR [14, 47-53]. Our main aim was to reproduce the helix 12 transformation from an unfolded conformation to an agonist or an antagonist form and to describe the transition path and the free energy differences between them. The aMD approach does not represent the real dynamics of the system, but is a valuable tool to identify the conformational states that exist in the free energy landscape and to explain which substructural changes lead to these states. Here we use as a starting point the extended H12 folding as solved in the pdb id 1a52 structure. We presume that this structure is an artifact that does not represent the real apo form of ERα, thus we named here this initial H12 conformational state as an unfolded conformation.

METHODS

Protein structures employed

The X-ray structure of pdb id: 1a52 was employed as an initial structure of the ERα apo state, which is named here as an unfolded state. It has been noted that this is a hypothetical crystal structure, because the extended H12 interacted with another monomer of two ERα dimers (i.e., tetramer) using an s-s bond [10]. This structure was employed in particular for two main reasons:

1) Most of the similar MD simulations performed so far, which studied H12 dynamics in ERα, had been based on this structure, and all had failed to sample significant H12 conformational changes and reproduce the agonist or antagonist states such as those seen in many X-ray structures; thus, making possible direct comparisons between the approaches used.

2) Instead of generating a random initial H12 position in our simulations, as done previously [14], or generating one that is close to LBD, which would obviously affect the initial conformational state, we selected the extended H12 conformation as seen in pdb id 1a52

crystal structure because it differs significantly from those observed in an agonist or an antagonist position; hence a very large energy barrier would have to be overcome so as to describe the whole conformational landscape. The latter fact makes the simulations much more reliable.

Initially, only the activation helix in monomer structures was investigated. For the starting antagonist H12 conformation (aMDm9-11 runs) a pdb id 3ERT (a complex with 4-hydroxytamoxifen) was chosen, whereas for an agonist helix 12 position (aMDm12-13 simulations) a pdb id 1QKU was selected (a complex with E2). These structures were preferred because they had been the most often used in similar studies and also there is no significant difference in H12 position between the hundreds of X-ray structures solved. Numerous initial structures were first prepared in order to study and reproduce the H12 transition from an unfolded state to either agonist or antagonist and an exchange between these conformations (see Tables S1 and S2 in the Supporting information). In particular, for the simulations aMDm9-11, 14 and aMDm24 the ligands were removed from the X-ray structures in an attempt to remove their suspected impact on the H12 conformation. The coactivators (when present) were also removed, except in simulation aMDm21. For simulation aMDm17 the H12 was extended with additional 7 residues, not solved in pdb id 1a52, which was constructed based on the helix 12 conformation as present in the pdb id 1QKU agonist structure.

Conventional Molecular dynamics (cMD)

Conventional molecular dynamics (cMD) were carried out using the Amber 12 suite of programs and the Amber99SB force field [54-56]. A truncated octahedral of TIP3P water molecules, 12Å dimension in each direction and counterions were added to obtain the final solvated system, which consisted of nearly 57 000 and 114 000 atoms for an ERα monomer and dimer, respectively. Initially, the systems were energy-minimized in two steps. First, only the water molecules and ions were minimized for 6000 steps while keeping the protein structure restricted by weak harmonic constrains of 2 kcal mol^{-1} $Å^{-2}$. Second, a 6000 step minimization with the conjugate gradient method (convergence criterion of 0.1 kcal mol^{-1} $Å^{-2}$) on the whole system was performed. Furthermore, the simulated systems were gradually heated from 0 to 310 K for 50 ps (NVT ensemble) and equilibrated for 3 ns (NPT ensemble). The production runs were performed at 310 K in a NPT ensemble. Temperature regulation was done using a Langevin thermostat with collision frequency of 2 ps^{-1}. The time step of the

simulations was 2 fs with a nonbonded cutoff of 9 Å using the SHAKE algorithm [57] and the particle-mesh Ewald method [58].

Accelerated Molecular dynamics (aMD)

The accelerated molecular dynamics (aMD) simulations provide the possibility to sample the conformational space much better and to detect the local energy minima that remain hidden in the cMD calculations [59-61]. aMD does not represent the real protein dynamics but is useful to investigate which helix 12 conformations are present on the free energy landscape and how they change in the different ERα forms, i.e. monomer or dimer. Moreover, aMD simulations typically boost the sampling by about 2000 times compared to cMD [55]. Thus, one can consider that the sampling performed by a 50 ns aMD trajectory might be equal to that of hundreds of nanoseconds of cMD simulation. aMD modifies the energy landscape by adding a boost potential $\Delta V(r)$ to the original potential energy surface when $V(r)$ is below a predefined energy level E:

$$\Delta V(r) = \begin{cases} 0, & V(r) \geq E \\ \dfrac{(E - V(r))^2}{\alpha + (E - V(r))} & V(r) < E \end{cases} \quad (1)$$

In principle, this approach also allows the correct canonical average of an observable calculated from configurations sampled on the modified potential energy surface to be fully recovered from the aMD simulations [55].

All of the aMD calculations were performed using the Amber 12 molecular modeling package [54-56]. In order to simultaneously enhance the sampling of the internal and diffusive degrees of freedom, a dual-boosting approach based on separate dihedral and total boost potentials was employed [55, 59-60]. This method may be compromised by the increased statistical noise but was also successfully applied in a similar studies [59]. The selections of the boost parameters E and α for the dihedral boost and the total boost were based on the corresponding average dihedral energy and total potential energy obtained from the above-performed cMD equilibration runs, respectively. The dihedral boost parameter, E_d, was set equal to the average dihedral energy obtained from the cMD simulation plus $N_{sr} \times 3.5$ kcal/mol, where N_{sr} is the number of solute residues; the α_d parameter was then set equal to $E_d/5$. For the total boost parameter, E_{tot}, the value was set to be equal to 0.16 N_{tot} plus the average total potential energy obtained from the cMD simulation, where N_{tot} is the total number of atoms; α_{tot} was simply set equal to 0.16 N_{tot}.

Simulation times between 50 and 100 ns were employed for the different protein configurations during our initial investigation on the ERα monomer, and the total length of the MD runs was over 1,2 μs (see Table S2). Further, to perform a more detailed analysis of the extended H12 conformation and those in a dimer, a number of new aMD runs with a length between 70 and 250 ns were performed (see Table S1). The simulations were performed at physiological conditions (about 310 K) instead of at 300 K, which give a slight additional boost of the aMD runs. Different temperatures were used only in three of the initial simulations in a monomer configuration – aMDd18-20 (see Table S2). The dual boost approach was implemented due to three main reasons. First, a set of aMD simulations executed on the similar PPARγ nuclear receptor showed that boosting only the dihedral term did not lead to reproduction of the known H12 agonist state when starting from an unfolded conformation (unpublished results). Second, based on the dozens of previous classical MD simulations and those initially performed in this study, it seems that when starting from an unfolded position, H12 initially drops into a very deep local minimum and the transition to its natural forms is difficult by the execution of regular cMD, which means that a sufficient boost should be applied in aMD runs.

Free Energy Calculations.

Free energy calculations were performed by the molecular mechanics Poisson−Boltzmann surface area (MM-PBSA) method using the MMPBSA.py script included in the AmberTools 14 package [54, 62]. The analysis was based on the most populated stated of the MD trajectories. In order to evaluate the individual contributions of the selected residues and/or structural elements, a decomposition of the free energy contributions approach was employed. Pairwise per-residue basis decomposition for the selected protein residues was chosen. Pairwise decomposition calculates the interaction energy between all pairs of residues in the system. The MMPBSA.py script was initially developed for calculating ligand−receptor interactions and currently requires the specification of a ligand. Thus, the protein was virtually divided into two parts: a ligand that consists of the helix 12 (residues 530−544) and a receptor for the remaining residues. However, all of the calculations were based only on summation of the pairwise free energies between the residues of interest taken from the output file. For instance, to access the interactions between the H12 and the remaining protein parts, we calculated all of the pairwise free energies between residues involved in these structural elements. All of the decomposed energies were calculated by the MM-PBSA method. The MM-PBSA is a widely used approach for free energy calculations, and a detailed description

of the methodology is available elsewhere [63-67]. Shortly, for each frame extracted from the MD trajectory (a 50 ps interval was set here), the enthalpic free energy was calculated as follows:

$$\Delta G_{bind} = \Delta G_{gas} + \Delta G_{solv} - T\Delta S \quad (2)$$

$$\Delta G_{gas} = \Delta G_{ele} + \Delta G_{vdw} \quad (3)$$

$$\Delta G_{solv} = \Delta G_{PB} + \Delta G_{nonpolar} \quad (4)$$

$$\Delta G_{nonpolar} = \gamma A + b \quad (5)$$

It included van der Waals (vdW) contributions, electrostatic contributions, as well as polar and nonpolar contributions to the solvation free energy (ΔGsolv). The gas-phase free energy (ΔG_{gas}) was obtained using the sander module of Amber 12, and the estimation of ΔG_{PB} was conducted by MMPBSA.py. $\Delta G_{nonpolar}$ was determined from eq 5, in which A is the solvent-accessible surface area, estimated using the Molsurf program (a part of the Amber 12 suite of programs) with a solvent probe radius of 1.4 Å, and γ and b are empirical constants set to 0.0072 kcal mol^{-1} $Å^{-2}$ and 0.92 kcal/mol, respectively. The entropy term (TΔS) was neglected in these calculations. As far as decomposition is concerned, a reasonable way involved splitting the electrostatic contribution to the free energy (ΔG_{PB}) into two parts, a self/desolvation energy and an interaction/pairwise energy. The self-energy is the value of ΔG_{PB} when only the residue of interest is charged while all other residues remain neutral; the interaction energy is the charging energy of the residue in the field of all other charged residues. Thus, the procedure applied here involves two numerical solutions of the Poisson−Boltzmann equation with differently charged states for each residue of interest.

Correlation Analysis

Cross-correlation maps of the residues' motion were used to identify the regions that were moving in or out of phase during the simulations. Thus, this method can provide information about the H12 conformation and how it reflects on the protein substructural changes and vice versa, thus specifying which structural elements are involved in the structural changes of the H12 position and the transition path, and what is their role, as well. The elements b_{ij} of the correlation matrix B were obtained from *ri* and *rj*, the position vectors of residues *i* and *j*, as follows:

$$bij = \frac{\langle (r_i - \langle r_i \rangle)(r_j - \langle r_j \rangle) \rangle}{\sqrt{(\langle r_i^2 \rangle - \langle r_i \rangle^2)(\langle r^2{}_j \rangle - \langle r_j \rangle^2)}} \quad (6)$$

Only the Cα displacements were considered during the calculations, which were performed using the Ptraj module of the AmberTools 14 program [54, 68]. All of the frames of the MD trajectories were aligned on the preequilibrated structure (the input for the aMD runs) with an interval of 50 ps. The analysis was based on the entire MD trajectories. The value of b_{ij} can vary from −1 (completely anticorrelated motion) to +1 (completely correlated motion). It should be noted that with the formula above it is impossible to detect orthogonal correlative motion.

Principal Component Analysis

PCA analysis is another helpful tool for studying the differences in the protein dynamics, in particular those of the substructural elements of interest. In short, PCA is a dimensional reduction technique that allows the isolation of the most significant conformational differences among a set of structures. In the MD analysis these structures were extracted as snapshots from the aligned MD trajectory, and then the correlation matrix was calculated and diagonalized, providing an orthogonal set of eigenvectors representing linearly independent modes of conformational changes. These are the principal components (PCs). The eigenvalue associated with each PC is a measure of the variance in the original data set explained by that component. In order to identify the conformational changes provoked by the H12 position and the substructural conformational changes and vice versa, the entire MD trajectories were included in these calculations. In this work, the PCAs were performed on the Cα coordinates using the Ptraj and Cpptraj modules of the AmberTools 14 [54] package, and the results are presented as heat maps of the first two PCs (PC1 and PC2), which were used as reaction coordinates for reweighing the aMD frames. Projections of the motion corresponding to PC1 onto the protein structures were calculated using the Prody software [69].

aMD simulations reweighting procedure

Reweighting of biased aMD frames is an important procedure and was performed using a previously described protocol [55, 70]. For an aMD simulation of a N_{atom} protein system, the probability distribution along a selected reaction coordinates *A(r)* is written as *p*(A)*, where r denotes the atomic positions $\{r_1,\cdots,r_N\}$. Given the boost potential *ΔV(r)* of each frame, *p*(A)* can be reweighted to recover the canonical ensemble distribution, *p (A)*, as:

$$p(A)_j = p^*(A_j)\frac{\langle e^{\beta\Delta V(r)}\rangle j}{\sum_{j=1}^{M}\langle e^{\beta\Delta V(r)}\rangle j}, \qquad j=1,...,M \tag{7}$$

where M is the number of bins, $\beta = 1/k_BT$ and $\langle e^{\beta\Delta V(r)} \rangle_j$ is the ensemble-averaged Boltzmann factor of $\Delta V(r)$ for simulation frames found in the j^{th} bin. The above equation provides an exponential average algorithm for the reweighting of aMD simulations. Further, the exponential term was approximated here by the summation of the Maclaurin series of boost potential $\Delta V(r)$ with the reweighting factor rewritten as:

$$\langle e^{\beta\Delta V} \rangle = \sum_{k=0}^{\infty} \frac{\beta^k}{k!} \langle \Delta V^k \rangle \quad (8)$$

where the subscript *j* has been suppressed. The Maclaurin series expansion to the 10th order has been used here. Note that the Boltzmann reweighting factors are often dominated by high boost potential frames [71-73]. In order to minimize this effect, these frames were identified in our output log files and removed from the trajectories composed for the analyses. This should reduce the noise but it should be noted that currently it seems that there is still no method that can fully eliminate the noise produced by the boost applied in the aMD simulations [70]. However, the reweighting procedure should not lead to a dramatic change in the main conformational state populations because no large boosts have been observed during the sampling of these main conformations and consequently affecting the main conclusions, although the populations of the intermediate states might be influenced by the introduced noise. A toolkit of Python scripts "PyReweighting" [70] was used to reweight the aMD simulations to calculate free energy profiles and also in the house scripts developed. In many cases, when the number of simulation frames within a bin was lower than a certain limit, i.e., cutoff, then the bin was not sufficiently sampled and thus excluded for reweighing [70]. To overcome this problem, the cutoff was determined by iteratively increasing it until the minimum position of the PMF profile does not change. The bin size was set to 500 and 700 for reweighting the aMD simulations here of a monomer and a dimer, respectively.

aMD convergence analyses

Recently there has been suggested a new and robust set of approaches for the intra and inter-simulation assessment of the MD simulations convergence, which were also applied in this work [74]. The "RMS average correlation" (RAC) functionality, which can be thought of as a pseudo-autocorrelation function for RMSD values, was used to access the convergences of the individual simulations. This essentially measures the convergence of the overall average structure at different time intervals within a single trajectory. For a given time interval or lag (τ) a straight coordinate running average over that time interval is performed over the entire

trajectory; each sliding averaged structure over the time interval τ is then either fit to the first averaged structure (time 0 − τ) or a reference structure specified by the user, and finally the average RMSD value of all averaged structures of length τ is calculated according to:

$$RAC(\tau)=\frac{\sum_{t=0}^{N} RMSD(AvgCrd(t,t+\tau))}{N-\tau+1} \quad (9)$$

where N is the total number of frames. At time τ=1, this is the standard average RMSD over the whole trajectory. When τ approaches the end of the trajectory length, the value approaches zero and loses meaning.

In order to assess the convergence of the internal motions between independent trajectories, we examined the overlap of histograms of the first two principal component projections (PC1 and PC2) [75-76]. First, to ensure that the eigenvectors obtained from each simulation being compared match, the coordinate covariance matrix (using only backbone atoms) was calculated using a combined trajectory from both simulations. Each frame of the trajectory was RMS-fit to the overall average coordinates in order to remove global rotational and translational motions. Next, the projection along these eigenvectors of each coordinate frame from the first simulation trajectory was calculated; this was then repeated for the second simulation trajectory. Finally, at each frame t a histogram for each simulation of the PC projection values for a given PC was constructed, and the overlap of these histograms was calculated using Kullback–Leibler divergence, KLD [75-76]:

$$KLD(t)=\sum_{i=o}^{M} \ln\left(\frac{hPC1_N(t,i)}{hPC2_N(t,i)}\right) hPC1_N(t,i) \quad (10)$$

where $hPCX_N(t,i)$ denotes bin *i* of the histogram from trajectory X for the projection of PC N using data from frames 0 to t, and M is the total number of histogram bins (400 in this case). The KDL methods histogram was also constructed using a Gaussian kernel density estimator with a bandwidth obtained via the normal distribution approximation of the PC data. The above analyses were carried out by AmberTools 14 package [54]. It should be noted that indeed other approaches for verification on the convergence are also possible [77-78].

Transition paths analyses

To illustrate the observed most likely transition paths from an unfolded to the well known experimentally states, the intermediate states were identified based on the best representative conformation between the most populated clusters observed between these main states, after reweighing the snapshots, whereas the transition points were the points between them as were

detected by the RMSD analysis. For each of the clusters the free energies were calculated and the short lived states formed by this intermediate binding were identified. Finally, the RAC analysis was also used for the transition events detection. Several movies were supplied to clearly illustrate the identified transition paths.

RESULTS AND DISCUSSION

cMD and aMD simulations of the monomer structure with H12 in an unfolded state

Tables S1 and S2 present summarized information about all performed simulations. Initially, we followed protocols that were similar to those employed previously [47-48]. Numerous independent aMD and cMD runs were performed on an ERα monomer, named here as aMDm9 to aMDm24 (see table S2), starting from either an agonist or an antagonist H12 conformation. However, no transition between these states was detected. Further, we used the assumed as an apo structure pdb id 1a52. A new MD simulations set was run again with different simulation conditions (see Table S2 for details). The water buffer, van der Waals cut-offs, temperatures and the presence of a ligand and/or a coactivator were varied as well. For instance, in an effort to achieve a better sampling, we used a simulated annealing approach and applied a high temperature of 800 K solely on the H12, holding the remaining part of the receptor unaffected by a harmonic constraint of 2 kcal.mol^{-1}.Å^{-1}. A detailed analysis of these simulations, such as the H12 starting conformations from either an agonist or an antagonist position, is not within the scope of this paper, but they were informative for our further investigation steps. In general, the result of all these 16 simulations with a total length of 1.1 μs was the same – no significant changes in the conformation and the transition of the helix 12 were observed in monomer configurations.

It can be argued that the MD runs above were repeated only up to two times and no general conclusions can be derived. Therefore, an additional set of 8 independent aMD simulations (aMDm1-aMDm8), starting from an unfolded state as seen in pdb id 1a52 X-ray structure, were performed (see Table S1). However, we were again not able to replicate an agonist state of the activation helix, starting from this unfolded H12 conformation. To test the sampling capability of the aMD approach in our concrete case and in particular to examine how the H12 motion affects the internal part of the LBD, we compared the principle component analysis (PCA) results for the receptor without helix 12 (residues 312 ÷ 530). The protein

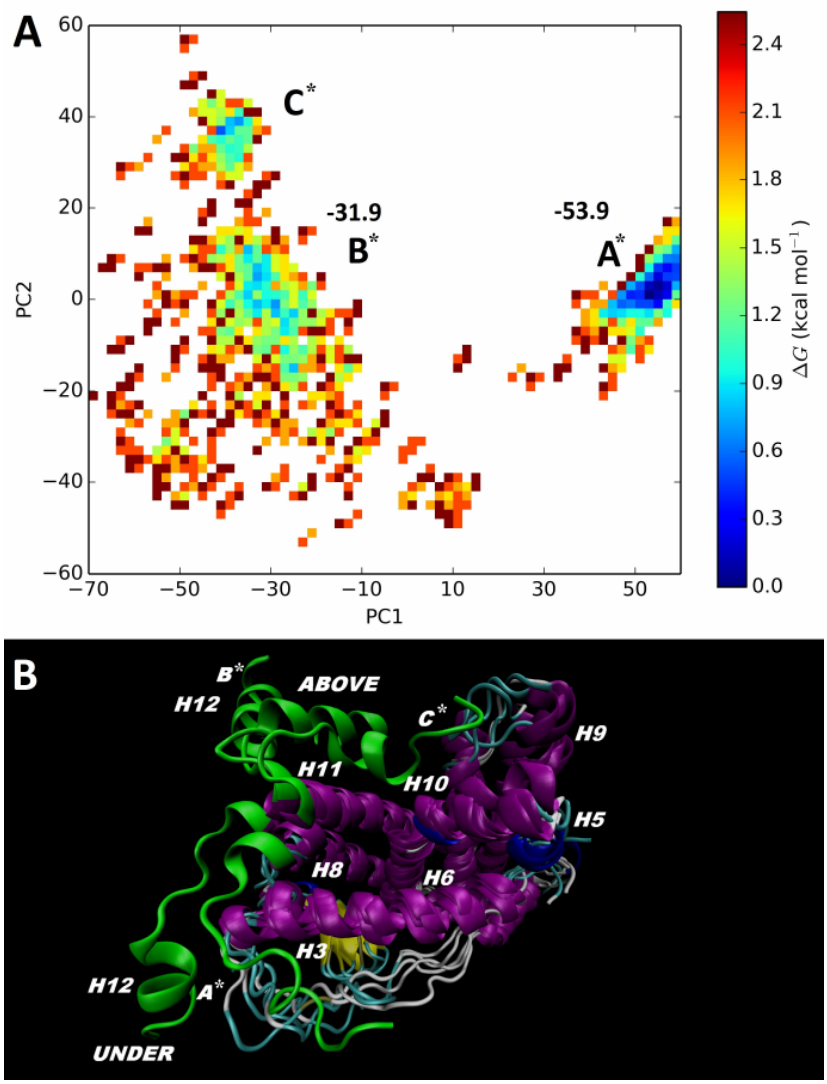


Figure 1. Detected helix 12 two main clusters in a monomer, which are perpendicular and parallel to the helix 11, and named here as an "under" and an "above" conformations, respectively. ***(A)*** PCA analysis of the aMDm2 trajectory, which describes the above observed positions, presented as a heat free energy map (kcal/mol), and ***(B)*** Clustering analysis of the most populated states of all performed aMD monomer runs, presented as aligned pdb structures. Calculated by MM/PBSA method, H12 entalpic free energies (kcal/mol) were also noted.

dynamics obtained by a 100 ns-long cMD (cMDm9) and 70 ns and 100 ns aMD runs (aMDm1 and aMDm3), respectively, were compared, which clearly showed that the accelerated dynamics sampled much better the ERα conformational space (see Figures S1A - S1C), thus supporting the conclusion that aMD is a valuable tool for conformational sampling. This was also indicated by a simple RMSD analysis (see Figures S2). However, these aMD simulations also showed a clear difference due to the fact that the aMDm1 explored the space under helix 11, where less structural transformations were observed, but the aMDm3 described several states from the much more flexible region above helix 11. To study this difference, a PCA analysis of solely the H12 motion was performed and showed

several low and highly populated states, i.e. long and short-lived H12 conformations, which were observed in almost all simulations, but their populations were different (see Figure 1A and movie 1). These conformations which had been captured by the cluster analysis performed were visualized as aligned pdb files (see Figure 1B). According to the combined data, one can split the conformational space into two main clusters with the helix 12 conformation, which is perpendicular and parallel to the helix 11, and we named these clusters here as an "under" and an "above" position, respectively. To illustrate these positions, the aMDm1 and the aMDm3 simulations were chosen as representative because they sampled in the best way the individual helix 12 populations mentioned above. In each of these clusters the H12 conformation varied, populating the space close to the H3/H11 and the H11 helices, respectively. The most populated H12 cluster, according to all the performed simulations, was the one placed between the H3/H11 helices (under the H11), but the close contact and the opposite motion of H3 did not allow the H12 to overcome this barrier and transit into the agonist conformation. The aMDm1, provided much better sampling of H12 than the cMDm9 but it was still close to H3 (see Figures S1B and S1C). In the case when H12 was above H11 it could not approach its agonist binding site due to the closure of the H11 and H5 helices. The mean distance between the Cα atoms of Met523 of H11 and Cys382 of H5 was only 8.5 Å, compared to 10.1 Å for a typical agonist structure (pdb id 1qku). In fact, the simple visual and root-means squarer fluctuation (RMSF) analyses (see Figure S3 and movie 1) showed that the portion of H11 (residues 525-531), i.e. the part just before H12, underwent some of the largest structural changes, and H11 moved up and down periodically. This movement was different in the different clusters, but led to the same effect, restricting H12 from adopting its known agonist conformation. What also contributed to this restriction was the collective motion in the AF-2 pocket, which can be revealed by more advanced analyses as presented in the next section (see details below). Note that both the RMSD and RMSF analyses represent only the structural changes detected by the aMD approach and not the real receptor dynamics.

The overall results from our initial aMD runs in a monomer did not show any transition between an unfolded conformation to either an agonist or an antagonist state, indicating the existence of very high energy barriers that are difficult to overcome. Identical results were previously obtained by all performed conventional MD studies which used simulation windows in the range of 5 to 30 ns, and recently up to 100 ns [14, 47-48]. These studies explored the same initial crystal structure (pdb id 1a52) as the one in the present work and the same H12 starting conformation as the one in this study, but were not able to reproduce the well-known from X-ray experimental H12 agonist or antagonist conformations. Similar

results have been obtained by set of MD simulations of other NRs, as well [50-53]. Indeed, one can suggest that the above mentioned cMDs were limited because of the short simulation time and therefore suffered from inadequate sampling and failed due to inability to overcome high potential energy barriers. However, in our study 8 independent dual-boosted aMD runs in a combination with several cMDs were performed with a total coverage time of 1.9 μs. These data have confirmed results obtained by earlier cMDs, thus indicating that the H12 transition from such an unfolded conformation to agonist or antagonist states is presumably not a sampling problem.

Correlation and PCA projection analyses of the aMDs in a monomer

Although we failed to reproduce the H12 transition to a proper agonist or antagonist form in a monomer, we analyzed identified clusters of the activation helix in the conformational landscape. Motivation for such an analysis also came from the previous MD studies which had detected a similar H12 motion. Indeed, it is interesting to answer the question why in so many independently performed molecular dynamics simulations the extended helix 12 form was not able to seat in the active site and moved in particular to either H3 or H11 helices in the monomer but did not approach the LBD. To reveal the formation of the above mentioned clusters and their role in the LBD reorganization, both correlation analysis and visualization of the motions by projections of the first principle component on the protein structure (in short PCA projections) were performed.

The aMD runs 1 and 3 were chosen as representative for the cluster with H12 with a position under and above H11, respectively. Correlation analysis identified that the motion between the pairs H3/H5, H5/H10, H8/H11 and notably between H3, H5/H6, H7/8, H11 and H12 helices had correlative motions, i.e. these helices moved simultaneously in a similar direction (see Figures 2A, 2B, S4A and S4B). This is also clearly seen from the projections of the first principle component representing the protein elements motions during these simulations (see Figures S5A and S5B). Further, the motion of the central part of H11 (residues 512-520) correlated with H5/H6 helices, creating a network of collective motions in LBD. Remarkably, most of the structurally important for E_2 binding residues in LBD, such as Asp353, Arg394 and His524, were positioned in regions with highly correlative motions, confirming that allosteric changes in LBD are critical for both E_2 binding and H12 helix conformation.

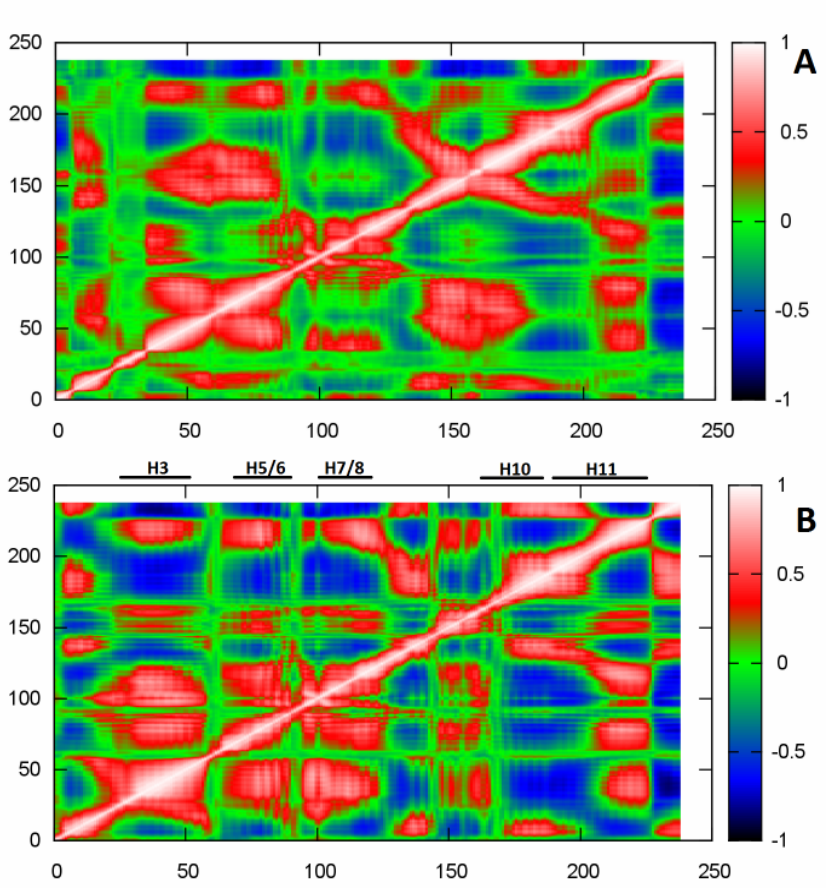


Figure 2. Observed correlations between the residues motion in ERα during simulations ***(A)*** monomer run 1 (aMDm1) and ***(B)*** monomer run 1 (aMDm3). The receptor residues were renumbered and start from 1 instead of 306. Thus some of the structurally important for estradiol (E_2) binding residues such as Asp353, Arg394 and His524 correspond to the Asp48, Arg89 and His219 on this correlation map, respectively. The red areas mark the residues with correlative motion, i.e. those which move simultaneously in the same direction and have a positive correlation coefficient, whereas the blue ones show those with an anticorelative motion, i.e. those which move in the opposite direction and have a negative correlation coefficient. The green areas are neutral.

The identified correlative motions shared a lot of similarities between the performed simulations, but there were also significant differences that might explain the observed H12 conformational changes. The main difference seen in the aMDm1 run compared to aMDm3 was the lack of significant H12 correlative motion with H7/H8 helices and a less noteworthy one with H3 (see Figure 2A). During the aMD1m run the distance between H8 and H12 was much larger than the one seen in the aMD3m run, hence the transition to the final state (perpendicular and close to H3) presumably does not affects the motion of above helices. Different motion was detected also between H5/H6 and H7/H8 helices. These results were also supported by the projection of the PC1, which showed that these structural elements moved in reverse directions compared to those in the aMD3m run (see Figures S5A and S5B). It should be noted that the H12 and H11 motions had a correlative motion in all these

independent aMD runs, which means that in both clusters under and above H11, the interactions with H11 are important for the H12 conformation.

Combined correlation and PCA projection analyses led to the conclusion that the helix 12 movement was determined by the highly correlative motions between the LBD structural elements and in particular the AF-2 region or vice versa, the motion of H12 provoked the observed conformational changes in LBD. However, the fact that the LBD structural elements move mostly in a different direction from that of H12 means that it is more likely that H12 alone is significant for the whole LBD motion in a monomer. When H12 adopts a conformation above H11 (aMDm3 run), then H3 and H5/H6 helices move highly correlatively in a left and down direction on the plane of Figure S5B, thus closing the natural H12 binding pocket. In addition, helix 11 pushes the activation helix away from the LBD. Correspondingly, H11 moves in an up direction. However, when H12 is in a position under H11 and in a close position to H3 (aMDm1 run), the above mentioned helices move exactly in the opposite direction, restricting H12 from overcoming the interactions with H3 and trapping it into the active site (see Figure S5A). Indeed, there were variations between the H12 conformations in individual aMD simulations, but the conclusion above is general for all the preformed runs. Notably, the fact that in the presence of the same ligand (E_2), the H3, H5 and H11 helices showed different motions for the observed two opposite H12 conformations, indicates that the ligand is not involved in and does not affect the H12 transition process. Indeed, this does not mean that the ligand is not involved in the H12 stabilization of the final H12 conformation of the natural ERα form.

aMD simulations of the dimer and transition of the H12 helix to an agonist and an antagonist form

Nine independent aMD runs were performed for an ERα dimer form (see Table S1). Surprisingly and in contrast to a monomer, in a dimer we were able to reproduce successfully both the agonist and antagonist H12 conformations. The activation helix of ERα was able to approach the "mouse trap" in 6 of totally 9 runs. Two different transition paths (Path 1 and Path 2) were detected (see movies 2 and 3). These paths were almost identical to those seen in a monomer, i.e. H12 can approach helix 11 "under" or "above". This simply means that similarly to the observed point mutations in the H12 and/or LBD pocket, the dimer interactions perturb the ERα LBD residue interactions (see below), allowing H12 to switch its conformation much easier. It also indicates that the dimer formation lowers the energy

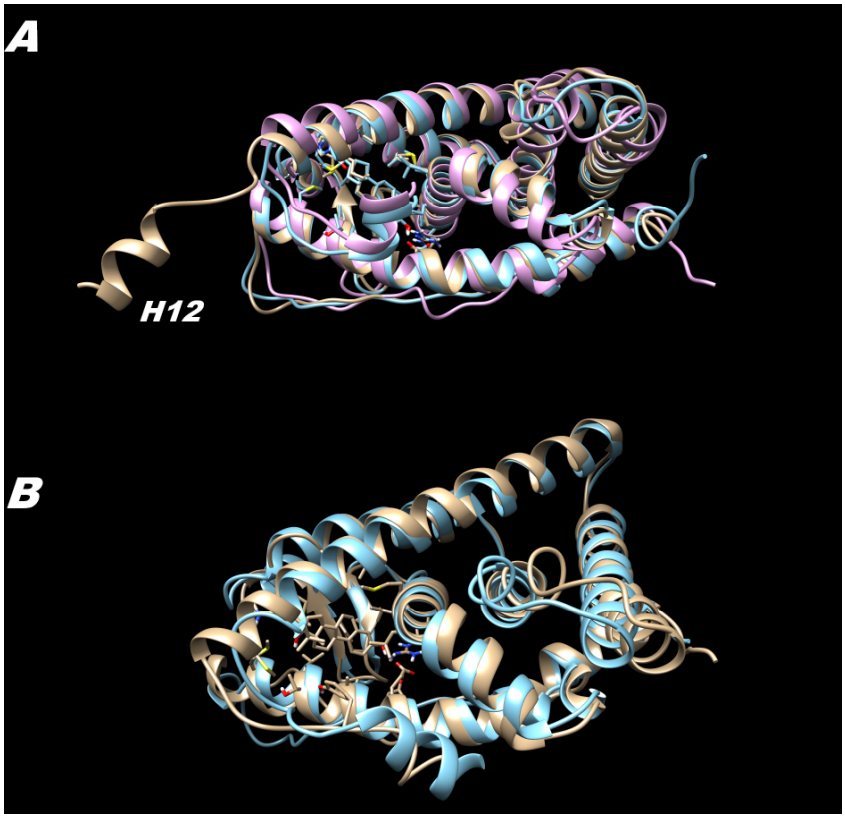


Figure 3. A comparison between: ***(A)*** the agonist conformation obtained via simulation aMD4 (in sky blue color), pdb id 1a52 (in wheat color) and pdb id 1qku (in pink) and ***(B)*** the antagonist position obtained via the same MD4 run (in wheat color), i.e. path 1 and X-ray structure pdb id 3ert.

barriers which H12 needs to overcome when approaching its agonist or presumably natural apo conformations. The best agreement with the well known from the X-ray data agonist state was achieved via the path "under" H11 helix (Path 1), i.e. the jumping over helix 3 close to helix 5 (see Figure 3A, movie 2). There was a remarkable overlap between experimentally solved conformations, as for example pdb id 1qku, and those obtained by our aMD simulations. The RMSD of the backbone atoms was equal to only 1.8 Å for aMD run number 4 (aMD4). Considering that the initial RMSD between H12 in the extended conformation, as seen in pdb id 1a52 crystal structure, and the agonist form is more than 25Å, it can be clearly concluded that this simulation precisely reproduced a transition path. The aMD1 run showed that the path exploring the conformational space "above" H11 (Path 2) seemed to be less populated, H12 was not properly folded and its conformation resembled to a less extent those seen in most agonist X-ray structures (see movie 3). Moreover, the agonist state via this path was also short lived and after 20 ns of simulation time transited to an antagonist-like conformation and then unfolded again.

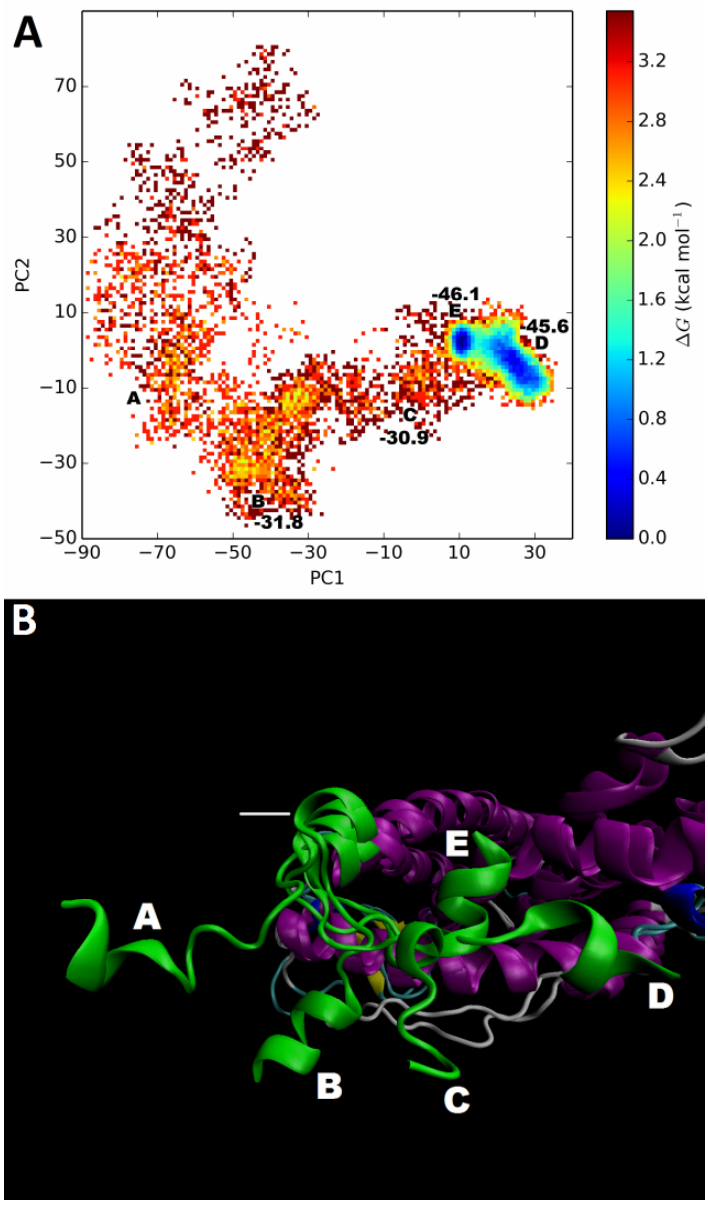


Figure 4. Transition of H12 via path 1 in a dimer based on ***(A)*** PCA analysis of the combined aMD2 and aMD4 trajectories, presented as a heat free energy map (kcal/mol), and ***(B)*** clustering analysis of the most populated states presented as aligned pdb structures. Position A is the unfolded conformation similar to X-ray pdb id 1a52, B and C represent the identified intermediates to antagonist (D) and agonist (E) states. Calculated by MM/PBSA method, H12 entalpic free energies (kcal/mol) and movement of the end part of H11 were also noted.

Notably, during the aMD runs 4, 5 and 7, which explored Path 1, we observed an initial H12 transition to a pure antagonist conformation that existed for about 10 to 20 ns, but then stabilized to an agonist form. In fact, the simulation which reproduced the H12 agonist conformation in the best way (aMD4 run) initially transited via the antagonist form too, but it was not fully folded (see movie 2). Moreover, simulation aMD2 reproduced only the antagonist helix 12 conformation, which was stable during almost the whole aMD run, and was properly folded compared to the known X-ray data (see Figure 3B and movie 4). Thus,

the aMD runs showed a stable antagonist conformation of helix 12 despite E_2 presence in the LBD. The RMSD value of 2.7 Å that was obtained by the comparison of the aMD2 results and antagonist X-ray data (pdb id 3ert) proved that the antagonist form was precisely reproduced too.

The reweighed and combined aMD2 and aMD4 PCA plots and the performed clustering analysis revealed that the agonist and antagonist states were the most populated, indicating that these conformations were the most likely to exist in ERα (see Figures 4A and 4B). All known H12 conformations were sampled well and many new ones were identified on the conformational landscape. Although the agonist and agonist clusters were almost equally populated, the position of the agonist cluster, i.e. the coordinates, was less variable than the antagonist one, indicating that during our simulations helix 12 was more flexible in an antagonist state. During the aMD4 run the agonist population was much better presented whereas in the aMD2 only the antagonist state was expressed. Several short-lived intermediate states were also presented on the free energy landscape, better expressed in aMD4 than in aMD2, populating mainly the space under H11 (i.e. close to H3), but the extended conformation was rare and not well sampled. During the preparation of the initial version of our manuscript, a coarse-grained (CG) model (not all atom-level of description), which recreated the switch between an agonist and an antagonist state, was published [79]. Although this work was focused on a specific mutation, D538G, and did not provide insight into the role of the LBD substructural elements, an identical result about the H12 agonist and antagonist states populations was obtained.

Thus, our results confirmed the recently suggested hypothesis that the pdb structure with an id 1a52 was presumably an artifact [14]. These data also clearly suggested that both the agonist and the antagonist states were possible in the ERα apo form because during our aMD simulations the E_2 ligand did not restrict the realization of the antagonist conformations, i.e. did not affect significantly the energies barriers. Hence, both the agonist and antagonist-like conformations were reproduced via the same transition paths, confirming that the constitution of an ERα homodimer favored the shift and that the presence of a ligand did not affect significantly the H12 transition process. This is also an indication that the agonist and antagonist H12 conformations are highly influenced by inter-dimer interactions (see below), and thus small perturbations, such as the point mutations and the type of ligands, can switch the H12 conformation from one preferred state to another. For instance, it has been reported that the H12 mutations such as L543A and L544A in the mouse ERα LBD changed

antagonists into agonists and the activation correlates with the activity of homodimerization [39].

Remarkably, the H12 transition to an agonist or an antagonist form was seen only in one of the dimer subunits. This was not an artifact from the initial conformation, because the transitions of H12 in individual simulations were observed randomly in both D1 and D2 dimer subunits. In the dimer subunit where no transition was observed H12 was seen to be close to or even jump over H3 during the simulation time, but was not capable to either properly fold, or occupy the expected agonist or antagonist position. Hence, coactivator peptide can bind to only one of the dimer units. This was an unexpected result, but it is in an agreement with obtained data by both biochemical and X-ray techniques [40]. In these studies it has been clearly shown that the full length (616-779aa) SRC-1 and SRC-2 coactivators bind only to one of the dimer units due to the allosteric conformational changes provoked in the second one.

Thus, according to our aMD simulations the presence of a ligand in LBD was not the leading factor for the H12 conformational changes and the transition from an extended to an agonist or an antagonist shape. This conclusion was supported by our monomer aMD runs, as discussed above. These results demonstrate that homodimer formation favors the constitution of an agonist and/or an antagonist conformation, and helps for the shift between the agonist and the antagonist states.

Visual, substructure distance, RMSF and RMSD analyses of the aMD's in dimer

To answer the question why such different results were obtained in our monomer and dimer aMD simulations initially, we performed visual, substructure distance, RMSF and RMSD analyses. From movie 2 (aMD4) and movie 4 (aMDm2) it is well visible that H11 moves in a different way in a dimer than in a monomer. In a monomer H11 moves up and down while in a dimer the pair of H11 helices move periodically close to and away from each other. Notably, during the transition between the intermediate states, the pair of H11 helices experienced the largest fluctuations opening the LBD pocket, assuring the H12 adoption to its agonist or antagonist conformation. This is also well visible from the RMSD and RMSF diagrams (Figures 5A and S3). Around the 4^{th}, 28^{th} and the 40^{th} nanoseconds of the aMD4 simulation, which best agreed with the experimental X-ray structures, large conformational changes in the last portion (the last 6 residues) of H11 in dimer subunit 1 (D1) were observed, which corresponded with and presumably constituted the observed intermediate states. It is

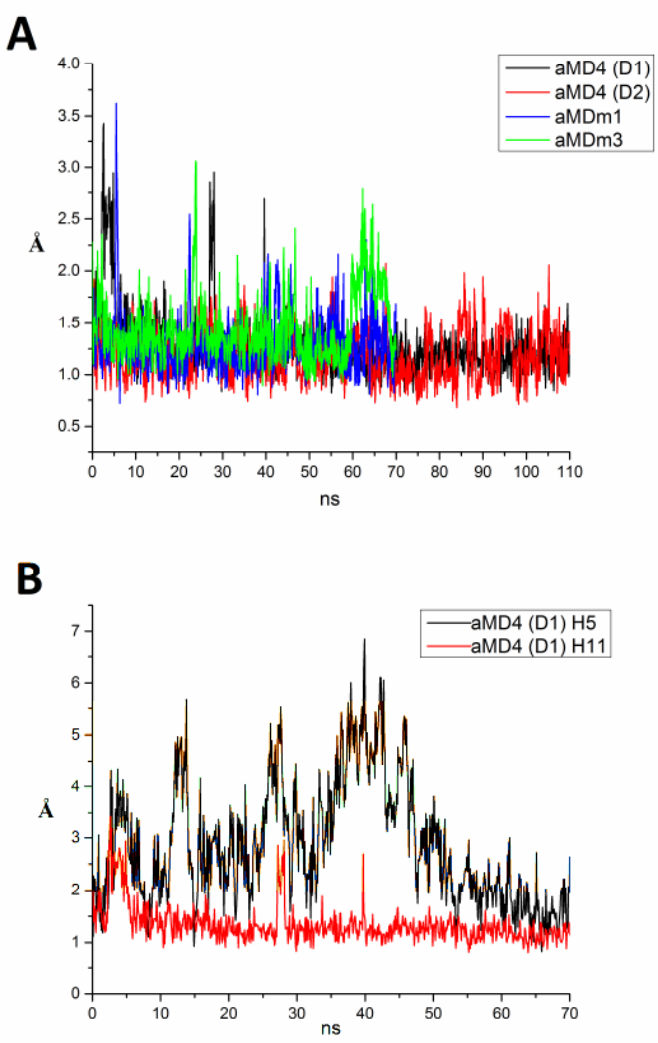


Figure 5. ***(A)*** Root mean square deviations (RMSD) of the last portion (the last 6 residues) of H11 obtained by simulations aMD4, aMDm1 and aMDm3. The dimer subunits 1 and 2 are denoted with D1 and D2, respectively. ***(B)*****.** RMSD of the H5 (372-381). Only the first 70 ns of the simulations are shown. Note that in both the H5 and H11 the larger fluctuations coincide (i.e. transition states).

noteworthy that the largest fluctuations in both H5 and H11 coincide (see Figure 5B), thus proving that correlative events occurred at the H12 binding pocket. Although the same portion of H11 in dimer subunits 2 (D2) looked stable during the whole simulation, when the H12 transition to an agonist state in D1 occurred, this portion of the complex became much more stable in the subunit D1. Compared to the monomer simulations aMDm1 and aMDm3, it became clear that only in a dimer H11 can be stabilized, while in a monomer it fluctuated randomly during the whole simulations, no matter whether they sampled a population under or above H11. The RMSF analysis (Figure S3) confirmed that even though the Cα atomic fluctuations were qualitatively identical in both homodimer subunits, their amplitude was much larger in those where H12 transition was achieved, i.e. dimer subunit D1. These structural differences compared to a monomer can be explained by specific interactions

mainly by two structural "hot" spots. First, a unique network of interactions was created between the central regions of the two H11 helices of the dimer in which the H8 of D2 and the H9/H10 of D1 were also joint. In these interactions there were involved Arg515, Asn519, Lys520 and Glu523 of H11 in D1, His513, His516, Asn519, Lys520 and Glu523 of H11 in D2, Glu423, Met427 and Arg434 of H8 in D2, and Glu385, Asn455 and Tyr459 of H10 and H9 in D1, respectively. It seems that Lys520 in both D1 and D2 plays a central role in these H-bonds network and during each H12 transition a reorganization in the network was observed, orienting the lysine in a different conformation towards the H8 and opposite the H11 residues. This also led to rearrangement of the pairs of Glu and Asn residues and they transformed their position too (see Figures S6 and S7). It is noteworthy that during the transition state to an antagonist position the structurally important for ligand binding residues His524 changed its conformation too and the stable H-bond with E_2 was completely destroyed. Further, the interactions of the H8 of D2 with the H9/H10 pair of D1 contributed to the conformation changes of the H5 in D1, thus affecting the whole LBD too. These were mainly hydrophobic interactions between the residues 459-463 of the H9 in D1 and 426-434 of the H8 in D2 when the Tyr459 of D1 was mainly involved. In particular, interactions between Tyr459 in D1 and His513 and Arg434 in D2 were identified and contributed for an agonist state to be established.

The changes in LBD in a dimer were mainly in H5 and the H9/H10 helices and they were mostly observed during the H12 transitions. In fact, this is the sole notable structural difference, except for the last part of H11 and the loops conformation, between the various intermediate states. Due to the strong interactions between helices 5 and 9/10, when H9/H10 move in some direction, then H5 translates in the same direction too. On the other hand, the close contact and interactions of the H9/H10 pair of D1 with the H7/H8 and the H11 of D2 create unique long range interactions that contribute to the H5 conformation in D1 and, consequently, to the LBD shape, and constitute the H12 transition and conformation. Indeed, such interactions would affect the coactivator binding, which interacts mainly with H5 too. When H12 is not placed in the LBD and is under H11, then H5 is in a closed position in respect to H11, but in the transition state it adopts an open conformation, allowing H12 to settle into an antagonist and then an agonist state (see Figure 5B).

The above findings might directly explain the high picomolar activity of some classes of agonists, such as TFMPV-E2, which affect the helix 7/8 conformation [80]. Only one small group attached to these strong agonists interacting with few additional residues affects the

helices mentioned, but a dramatic change in the activity has been observed. The last fact is difficult to explain by some stronger binding to the LBD [81]. According to our results, the extraordinary agonist activity of these inhibitors can be explained by affecting the whole dimer complex via stabilization of interactions between the above-mentioned H8/H11 residues and, consequently, the shape of the LBD and/or by stronger dimer subunits binding due to the changed helix 7/8 conformation.

Correlation and PCA projection analyses of the aMD's in a dimer

In order to explain in more detail the aMD results above, a correlation and a PCA projections analyses were performed. They were based on the aMD4 simulation where the transition of helix 12 to an agonist state was in the best agreement with experimental X-ray data. The observed correlations were plotted and the corresponding structural elements that were involved in the correlative motions and hence were likely to contribute to the H12 transitions were visualized and identified (see Figure 6). In general, the correlations in D1 and D2 were the same as in a monomer (aMDm1-path 1 too, see Figure 2A), but there was an important difference – no significant correlation between most LBD substructural elements, i.e. H3, H5/H6, H10 and H11, identified in monomer and the H12 motion was observed. This fact directly indicates that other factors are responsible for the H12 conformation in a dimer, which differ from those in a monomer. Three notable dynamical events between the dimer substructure elements were identified by the correlation analysis:

1) The motion of the second half of H11 and the whole H12 (residues 515-544) in dimer unit 1 (D1), where the H12 transition to an agonist and an antagonist states was observed, strongly correlate with the motion of the H9 of the dimer unit 2 (D2) (see Figure 6). The correlation was positive, indicating that these elements moved in the same direction. Further, the same portion of the H11 in D1 had an anticorrelative motion (i.e. in an opposite direction) with the second half of the H11 in D2 (residues 513-527). Finally, the motion of the first half of the H11 (residues 498-515) in D1 correlated with that of the whole H11 of D2.

2) The H5 and H9 helices of D1 had an anticorrelative motion with the H7 and H8 helices of D2. At the same time, H5 and H9 in D1 were highly correlative, proving that these helices moved in the same direction.

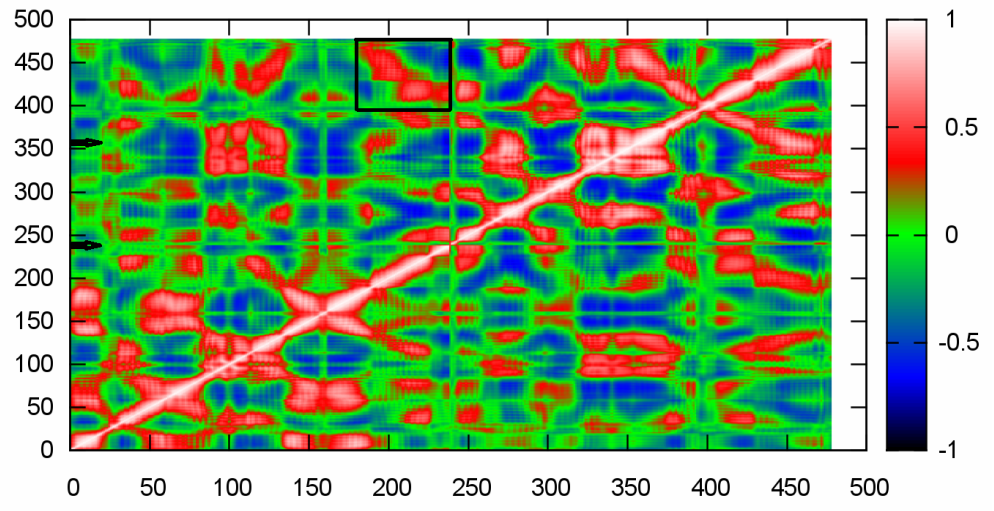


Figure 6. Observed correlations between the residues motion in an ERα dimer during simulations aMD4. The receptor residues were renumbered and start from 1 instead of 305 for subunit D1 and from 240 to 480 for D2. The red areas mark the residues with a correlative motion, i.e. those which move simultaneously in the same direction and have a positive correlation coefficient, whereas the blue ones show those with an anticorelative motion, i.e. those which move in the opposite direction and have a negative correlation coefficient. The green areas are neutral. The discussed and significant for the H12 conformation substructural movements are indicated with arrows and a square. The H11 anticorelative motion was centered at residues 217 in D1 and 456 in D2 (Met522), respectively.

3) The H6 and H7/8 helices of D1 had a correlative motion with the H7 and H8 helices of D2.

These correlations confirm the simple visual analysis of the events sampled by the aMD approach. They also agree with the conclusion that the long range interactions significantly alter the dynamics of the LDBs in other NR such as PRX homodimer [82]. More details can be achieved by PCA components projections analysis visualizing the directions of the motions and thus showing how the correlative motions detected above contribute to conformational alterations and why the changes in the ERα dimer are different in a monomer (see Figure 7). In agreement with the correlation analysis, the H11s of both D1 and D2 homodimer units move in different directions in a parallel to the common plane. The same conclusion can be drawn for the H8 of D2 and H5/H9 of D1, whereas the H5 and H9 in D1 move in the same direction. All structural elements in D1 constituting the LBD move in a different direction than those in a monomer, although H12 approaches the binding site via the same path – path 1 (see Figures S5A and 7). In a dimer, the H5/H6, H11 and H3 open the LBD, whereas in a monomer they contribute to its closure. In particular, the H5/H6 pair recedes from H11, allowing H12 to adopt an antagonist conformation and then transit to its agonist form. Thus,

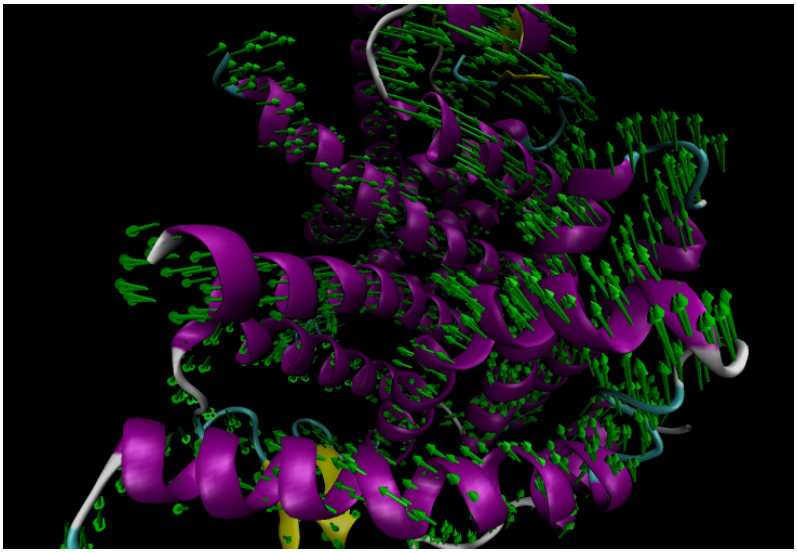


Figure 7. Observed protein motions represented by the projections of the 1st principle component (PC1) onto the protein structure of a dimer obtained by aMD4.

as was described and in support to the results in the previous paragraph, the unique direct and long-range interactions create a dimer motions network which contributes to the overall conformation of the ERα complex, affecting the motion in individual dimer subunits and contributing to the H12 transitions from one intermediate state to another.

Free energy analysis

Free energy analysis revealed why in a monomer the cluster of H12 conformations under H11 is the most populated. The path above H11 showed three populations with an entalpic free energy of -13.7, -18.8 and -31.9 kacl/mol, respectively. The entropy part was not included in these calculations. Conversely, the calculated free energy for the detected two populations under H11 had energies of -29.1 and -53.9 kcal/mol, which proved that this position of H12, when the starting point is its unfolded conformation, was much more energetically favorable and thus explained why it was the most populated. The H12 binding via the path above H11 were predominated by electrostatic interactions, whereas the Van der Waals ones via the alternative path, i.e. under H11, were major. It is noteworthy that the activation helix reaches such deep local minima in a monomer which can explain why it was not possible to jump over helix 3 and to be trapped into the LBD.

In a dimer the path above H11 (Path 2 - aMD1) was proved to be not energetically favorable with an energy of H12 of -29.6 kcal/mol. For Path 1 we calculated the free energies in 5 different positions representing the conformational change of H12 (see Figure 6). They can be

summarized as follows: H12 under H11 and close to H3, an intermediate state to an antagonist conformation, an antagonist position, an intermediate state to agonist and finally an agonist position. The measured free energies of this conformation were -31.8, -30.9, -45.6, -43.3 and -46.1 kcal/mol, respectively. It is notable that the free energies of the antagonist, agonist and the transit between them were very similar, proving that the switch between these conformations can be achieved by even small perturbations such as point mutations or interactions with a different type of ligands. It is also noteworthy that significant energy increasing was observed for H12 in an under H11 position in a dimer, compared to in a monomer, contributing to the easier switch from an unfolded to an active state in a dimer. Transitions via path 1 were also observed during aMD run 2 (aMD2), which sampled only the antagonist H12 position. The measured free energy was identical to those registered by the aMD4 simulation: -43.1 kcal/mol. During the observed transitions, the electrostatic and the Van der Waals interactions were with an equal contribution, but after the constitution of a stable agonist or antagonist state the Van der Waals and solvation H12 binding energies were much more significant than the electrostatic ones. Remarkably, during the initial transition to antagonist forms, the energy contribution came from only the first seven residues (residues 534-540) of H12 in all simulations, which practically controlled the whole transition process (data not shown).

A novel ER activity modulation model and structural basis of HDX and fluorescence data

All the conclusions from our simulations agree well with fluorescence and HDX experiments [34-36]. What is more, our results provide a structural basis of the notable HDX observation that the dynamics of helix 12 exhibit no significant stabilization on the binding of ligands, including E_2 [35-36]. In other words, although the position of H12 is crucial in determining the agonism versus antagonism, repositioning of this helix on ligand binding in the absence of a coactivator does not alter the dynamics of H12, which exhibits a rapid exchange in HDX experiments, regardless of the nature of the ligand [35]. We demonstrated here, and described in detail, that at least for E_2 the H12 exhibited equally populated agonist and antagonist states and its conformation was at least not directly ligand-dependent but was controlled by the observed inter-dimer long range interactions discussed above. Hence, a fast H12 switch between these states is also possible. Although the aMD approach does not represent the real dynamics, but samples the likely conformations instead, our results show that the 17β-estradiol presence cannot dramatically reshape the conformational landscape in

the manner which we observed between monomer and dimer configurations. Further, the observed here fluctuations of the last part of H11, followed by a deep stabilization after the H12 switch to an agonist state, agree with intrinsic tryptophan fluorescence experiments [34], revealing on a molecular level why this portion of H11 underwent significant changes during the H12 transition.

The importance of dimerization on the H12 antagonist conformation was recently experimentally confirmed [39, 83]. The dislocation of H12 caused ligand dependent LBD homodimerization involving the F-domain, the adjoining region of H12, thus indirectly indicating that dimerization process is necessary for the proper H12 antagonist folding. Our results confirmed this process, extended it also for the agonist state and provided structural and dynamical insight into the relationship between dimer formation and H12 conformation.

Previously preformed MD simulations on an ERα dimer suggested that dimerization can also play an important, and hitherto unexpected, role for the dynamics and regulation of ligand dissociation rates and binding mode throughout the nuclear receptor family [84-86]. Moreover, it was demonstrated that the perturbation of H12 in ERα^{Y537S} mutant slowed the hormone association rate (k_{+1}) by a factor of approximately 6 times and slowed the dissociation rate (k_{-1}) about 3-fold [87], and that agonists displayed much faster association rates than antagonists in the ERα wild type (nearly 500-fold difference) [88]. On the other hand, ligand binding strengthened the dimer complex *in vitro* and antagonists led to stronger interaction between dimer subunits [34, 38]. In addition, HDX experiments directly demonstrated that the dimer interface was selectively affected by ligand binding [34]. For instance, the binding of 4-hydroxytamoxifen-like ligands stabilized H11 but induced destabilization on the loop between the H9/H10 helices, thus affecting the interactions between the dimer units. Notably, it was also shown that there was a difference between the escape paths of the agonists and the antagonists, which can be linked to the observed kinetics, and it was concluded that the residues Lys520 and Glu523, located at the dimer interface in H11 and identified by our study as important for the H12 transitions, were also significant for the ligands binding path [84]. Thus, it seems that a common mechanism regulates both the H12 conformational state and the ligands binding.

Taking together all the above data and our results, we suggest a novel ERα activation/deactivation mechanism and, thereupon, the sequence of the binding events during receptor activity modulation. As described above, many strong antagonists and agonists do not significantly perturb directly the H12 conformation, as was widely thought till now, but, most likely, they alter the substructural elements involved in the interactions between the

dimer units. According to our findings, they might cause a change of the H12 equilibrium and the populations of its agonist and antagonist states, as well. Thus, we suggest that the dimer formation controls H12 conformational stability. After the dimer constitution, as shown here, equilibrium between the H12 agonist and antagonist states is achieved and the H12 conformation can undergo fast switches, in congruence with HDX experiments. However, when the hormone (i.e. agonists, antagonists, SERMs and etc.) accommodates to the LDB, it either stabilizes or destabilizes different substructural receptor parts which have been identified here as important for the long range inter-dimer motions and therefore for the final H12 agonist and antagonist states equilibrium. According to HDX experiments, these were H3, H5/H6, H8, H11, H9/H10 loop and the β1/ β2-sheets positioned close to H3, in agreement with our results [34]. This provokes both the dimer interface interactions and, presumably, in some cases direct H12 perturbations which can alter the activation helix equilibrium and reshape the conformational landscape, similar to several point mutations. This conclusion is also supported by the *in vitro* experiments which registered a much more stable dimer formation when both an agonist and an antagonist ligand are bound [38]. Indeed, this mechanism would be ligand-dependant and might change the H12 equilibrium to either an agonist or an antagonist state population. For instance, we concluded here that the E_2 did not affect significantly the H12 position and binding, and the activation helix was almost equally populated in the agonist and the antagonist states. However, the cluster of agonists conformation was less variable, thus indicating that even though the E_2 changed neither the inter-dimer interactions significantly, nor the H12 state directly, the perturbation provoked by the ligand in the helices identified above resulted in an overall receptor and thereby homodimer stabilization, as was shown by NMR results [32], and, finally, in an agonist state, as seen by X-ray technique under experimental non-physiologic conditions. Such a model can also explain the much faster agonist association rate compared to the usually bigger antagonist compounds that supposedly contribute to larger LDB transformations changes in the inter-dimer interactions [88]. Currently, an ongoing research is performed with the aim of specifying how the diverse set of agonist and antagonist ligands perturb the H12 conformation via the interactions between dimer subunits.

Finally, our results can be used in the design of new classes of ERα and, presumably, other NR modulators targeting dimer interface.

aMD convergence analyses

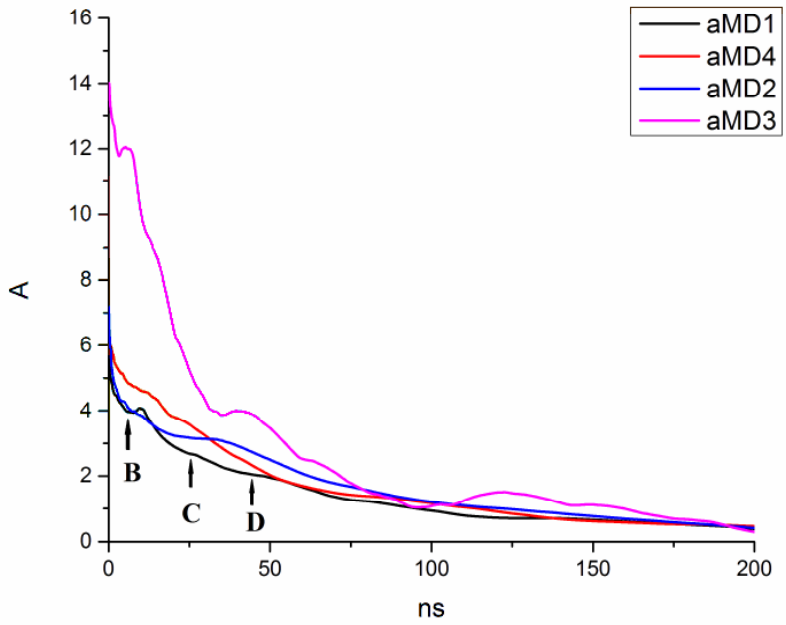


Figure 8. RMS average correlation (RAC) computed at different time intervals of the backbone atoms of running average structures over the full trajectories, with frames spaced at 100 ps intervals, calculated at each time interval from 100 ps to 250 ns, with an offset of 100 frames, referenced in the RMS fit to the average structure over the entire trajectory. The aMD1-4 runs was showed up to 200 ns for better representation of the plateau observed after 130 ns of the simulation time. The arrows indicate the observed in aMD4 run intermediate states detected as momentary plateaus, which correspond to those in Figure 6, and the transition states registered as climbs.

Obtaining a well-converged ensemble of structures is a key challenge within molecular dynamics simulation approaches, where "well-converged" implies adequate sampling of the conformational space of the system of interest. Enhanced sampling techniques are often applied to provide faster convergence than what is possible with traditional molecular dynamics, but this should be carefully examined within the performed study.

The convergence analyses of the individual trajectories were carried out by the RMS average correlation (RAC) method, which was recently introduced [74]. This is an advanced and rigorous approach for such an analysis. The aMD1, 2, 3 and 4 simulations were chosen for the RAC analysis because they represent both transition paths 1 and 2, reproduce in the best way the experimental data and were also mainly used for the analyses in the current study. Figure 8 shows the RAC results for aMD1, 2, 3 and 4 simulations, and indicates that a reasonable convergence for the above aMD runs after 60-70 ns of simulation time was observed, as well as a really good convergence after 130 ns. Thus, the simulations performed were well converged, ensuring the reliability of the results obtained. Remarkably, the RAC analysis provided also the opportunity for an additional transition states identification, and both the

above detected intermediate states (B, C, D from Figure 4) and the transition points in aMD4 run were well visible as climbs and plateaus, respectively (see Figures 8).

While the RAC is a measure of structural convergence within a single simulation, it is also of interest to measure how well two independent MD simulations converge with respect to each other. The convergence of the dynamic properties of the aMD4 and aMD5 simulations, which both led to an agonist conformation via the same path, was quantified by measuring the overlap of the PC histograms via Kullback–Leibler divergence (KDL) [75-76] as a function of the simulation time (see Figure S8). The PCA was performed on backbone residues and the first two PCs were relatively well-converged within only 35 ns. Next, an analysis of the overlap of the PC histogram from PC analysis in the Cartesian space, calculated from the combined aMD4 and aMD5 simulation trajectories, with an independent projection of the PCs on the separate trajectories, was performed (Figure S9). Identical distributions were observed, and the detected two peaks were only a bit shifted in this diagram indicating a reasonable convergence even between the performed simulations. Indeed, no convergence was observed between the simulations describing path 1 and 2, respectively. This is an indication that the execution of multiple aMD runs is highly recommended for systems with high energy barriers.

Using the aMD approach one can expect continuous jumping from one energy minimum to another to be observed during the whole simulation time due to the reduced barrier highs. However, for larger systems with significantly high energy barriers, such as H12 transitions in an ERα dimer, it seems that due to the presence of many deep local minima, a single aMD simulation gets into one of them after an indefinite simulation time. Our study showed that for both the ERα monomer and dimer multiple aMD should be run to explore adequately the conformational space.

CONCLUSIONS

In this paper we have investigated the likely conformations of the activation helix (helix 12) of Estrogen receptor alpha. A number of accelerated molecular dynamics runs on both ERα monomer and homodimer with a total length of 2.2 μs have been performed. Each of these runs were with lengths between 70 and 250 ns. The length of all accumulated simulations by the different approaches used reached about 3.3 μs. To our knowledge, these are the longest MD runs on a nuclear receptor reported to date. Based on these simulations, we were able, for the first time, to sample well the conformational landscape of the most important for the activity for all Nuclear Receptors helix 12, starting from an unfolded state, as well as to

reproduce precisely both the agonist and the antagonist conformations and to describe the transition paths between them even in the presence of an agonist ligand. The populations of these well known from X-ray data and many newly identified intermediate states were accessed and the free energy of binding for each of these H12 conformational clusters was calculated. All of the previously performed MD simulations had failed to recreate these states and to describe the transition paths between them. Further, the receptor substructural and inter-dimer interactions which lead to the H12 conformational changes were revealed. Based on these data, we have concluded that the dimer formation reshapes the H12 conformational landscape and is a leading factor for the activation helix conformation and dynamic. This not only agrees with experimental data but also provides a structural basis for several remarkable observations. Having combined previous data and our results, we also presumed that the ligands do not directly determinate the helix 12 position but rather stabilize/destabilize substructural parts of the receptor, which in turn affects dimerization and H12 conformation. Further research in this direction should be performed in order to specify the process above, but our results here have demonstrated that the presence of estradiol does not affect significantly the H12 conformation, as was observed in presence of the second dimer unit. Finally, we have proposed a novel ERα activation/deactivation pathway and, thereupon, the sequence of the binding events during receptor activity modulation, which helps for the explanation of many ERα dynamic features and the activity of some known ligands, and also provides a basis for the design of new classes of nuclear receptor modulators.

ACKNOWLEDGMENTS

Thanks to Maya Marinova for editing the manuscript and to Prof. Ilza Pajeva and Elina Mihaylova for their valuable remarks.

REFERENCES

[1] P. Germain, L. Altucci, W. Bourguet, C. Rochette-Egly and H. Gronemeyer, Nuclear receptor superfamily: Principles of signaling, *Pure Appl. Chem.*, 2003, **75**, 1619-1664.

[2] Nuclear Receptors Nomenclature Committee. A unified nomenclature system for the nuclear receptor superfamily, *Cell*, 1999, **97**, 161-163.

[3] H. Gronemeyer, J. A. Gustafsson and V. Laudet, Principles for modulation of the nuclear receptor superfamily, *Nat. Rev. Drug Discov.*, 2004, **3**, 950-964.

[4] J. T. Moore, J. L. Collins and K. H. Pearce, The nuclear receptor superfamily and drug discovery, *Chem. Med. Chem.*, 2006, **1**, 504-523.

[5] A. Berkenstam and J. A. Gustafsson, Nuclear receptors and their relevance to diseases related to lipid metabolism, *Curr. Opin. Pharmacol.*, 2005, **5**, 171-176.

[6] Ö. Wrange and J. A. Gustafsson, Separation of the hormone- and DNA-binding sites of the hepatic glucocorticoid receptor by means of proteolysis, *J. Biol. Chem.*, 1978, **253**, 856-865.

[7] W. Bourguet, P. Germain and H. Gronemeyer, Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications, *Trends Pharmacol. Sci.*, 2000, **21**, 381-388.

[8] U. Egner, N. Heinrich, M. Ruff, M. Gangloff, A. Mueller-Fahrnow and J. M. Wurtz, Different ligands-different receptor conformations: Modeling of the hERalpha LBD in complex with agonists and antagonists, *Med. Res. Rev.*, 2001, **21**, 523-539.

[9] R. T. Nolte, G. B. Wisely, S. Westin, J. E. Cobbs, M. H. Lambert, R. Kurokawa, M. G. Rosenfeld, T. M. Willson, C. K. Glass and M. V. Milburn, Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma, *Nature*, 1998, **395**, 137-143.

[10] D. M. Tanenbaum, Y. Wang, S. P. Williams and P. B. Sigler, Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 5998-6003.

[11] W. Bourguet, M. Ruff, P. Chambon, H. Gronemeyer and D. Moras, Crystal structure of the ligand binding domain of the human nuclear receptor RXRR, *Nature*, 1995, **375**, 377-382.

[12] J. P. Renaud, N. Rochel, M. Ruff, V. Vivat, P. Chambon, H. Gronemeyer and D. Moras, Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid, *Nature*, 1995, **378**, 681-689.

[13] C. Watanabe, H. Watanabe and S. Tanaka, An interpretation of positional displacement of the helix12 in nuclear receptors: preexistent swing-up motion triggered by ligand binding, *Biochim. Biophys. Acta*, 2010, **9**, 1832-1840.

[14] M. R. Batista and L. Martínez, Dynamics of nuclear receptor Helix-12 switch of transcription activation by modeling time-resolved fluorescence anisotropy decays, *Biophys. J.*, 2013, **7**, 1670-1680.

[15] M. G. Parker and R. White, Nuclear receptors spring into action, *Nat. Struct. Biol*., 1996, **3**, 113–115.

[16] M. Tanenbaum, Y. Wsng, S. P. Williams and P. B. Sigler, Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains, *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 5998- 6003.

[17] A. Wärnmark, E. Treuter, J.-Å. Gustafsson, R. E. Hubbard, A. M. Brzozowski and A. C. W. Pike, Interaction of transcriptional intermediary factor 2 nuclear receptor box peptides with the coactivator binding site of estrogen receptor α, *J. Biol. Chem.*, 2002, **277**, 21862–21868.

[18] A. K. Shiau, D. Barstad, P. M. Loria, L. Cheng, P. J. Kushner, D. A. Agard and G. L. Greene, The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen, *Cell*, 1998, **95**, 927-937.

[19] J. P. Renaud, N. Rochel, M. Ruff, V. Vivat, P. Chambon, H. Gronemeyer and D. Moras, Crystal structure of the RXR-ligand-binding domain bound to all-trans retinoic acid, *Nature*, 1995, **378**, 681–689.

[20] R. L. Wagner, J. W. Apriletti, M. E. McGrath, B. L. West, J. D. Baxter and R. J. Fletterick, A structural role for hormone in the thyroid hormone receptor, *Nature*, 1995, **378**, 690–697.

[21] L. Cantin, F. Faucher, J.-F. Couture, K. P. de J.-Tran, P. Legrand, L. C. Ciobanu, Y. Fréchette, R. Labrecque, S. M. Singh, F. Labrie and R. Breton, Structural characterization of the human androgen receptor ligand-binding domain complexed with EM5744, a rationally designed steroidal ligand bearing a bulky chain directed toward helix 12, *J. Biol. Chem.*, 2007, **282**, 30910–30919.

[22] C. Sonnenschein and A. M. Soto, An Updated Review of Environmental Estrogen and Androgen Mimics and Antagonists, *J. Steroid Chem. Mol. Biol.*, 1998, **65**, 143–150.

[23] L. Celik, J. D. D. Lund and B. Schiøtt, Exploring Interactions of Endocrine-Disrupting Compounds with Different Conformations of the Human Estrogen Receptor α Ligand Binding Domain: A Molecular Docking Study, *Chem. Res. Toxicol.*, 2008, **21**, 2195–2206.

[24] A. C. W. Pike, A. M. Brzozowski, R. E. Hubbard, T. Bonn, A. G. Thorsell, O. Engström, J. Ljunggren, J.-Å. Gustafsson and M. Carlquist, Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist, *EMBO J.*, 1999, **18**, 4608–4618.

[25] A. C. W. Pike, A. M. Brzozowski, J. Walton, R. E. Hubbard, T. Bonn, J.-Å. Gustafsson and M. Carlquist, Structural aspects of agonism and antagonism in the oestrogen receptor, *Biochem. Soc. Trans.*, 2000, **28**, 396–400.

[26] J. B. Bruning, M. J. Chalmers, S. Prasad, S. A. Busby, T. M. Kamenecka, Y. He, K. W. Nettles and P. R. Griffin, Partial agonists activate PPARgamma using a helix 12 independent mechanism, *Structure*, 2007, **15**, 1258–1271.

[27] B. A. Johnson, E. M. Wilson, Y. Li, D. E. Moller, R. G. Smith and G. Zhou, Ligand-induced stabilization of PPARgamma monitored by NMR spectroscopy: implications for nuclear receptor activation, *J. Mol. Biol.*, 2000, **298**, 187–194.

[28] P. Cronet, J. F. Petersen, R. Folmer, N. Blomberg, K. Sjöblom, U. Karlsson, E. L. Lindstedt and K. Bamberg, Structure of the PPARalpha and gamma ligand binding domain in complex with AZ 242; ligand selectivity and agonist activation in the PPAR family, *Structure*, 2001, **9**, 699–706.

[29] J. Lu, D. P. Cistola and E. Li, Analysis of ligand binding and protein dynamics of human retinoid X receptor alpha ligand-binding domain by nuclear magnetic resonance, *Biochemistry*, 2006, **45**, 1629–1639.

[30] B. C. Kallenberger, J. D. Love, V. K. Chatterjee and J. W. Schwabe, A dynamic mechanism of nuclear receptor activation and its perturbation in a human disease, *Nat. Struct. Biol.*, 2003, **10**, 136–140.

[31] T. S. Hughes, M. J. Chalmers, S. Novick, D. S. Kuruvilla, M. R. Chang, T. M. Kamenecka, M. Rance, B. A. Johnson, T. P. Burris, P. R. Griffin and D. J. Kojetin, Ligand and receptor dynamics contribute to the mechanism of graded PPARγ agonism, *Structure*, 2012, **20**, 139-150.

[32] V. Paramanik and M. K. Thakur, NMR analysis reveals 17β-estradiol induced conformational change in ERβ ligand binding domain expressed in E. coli, *Mol. Biol. Rep.*, 2011, **38**, 4657-4661.

[33] L. A. Luck, J. L. Barse, A. M. Luck and C. H. Peck, Conformational changes in the human estrogen receptor observed by (19) F NMR, *Biochem. Biophys. Res. Commun.*, 2000, **3**, 988-991.

[34] A. Tamrazi, K. E. Carlson and J. A. Katzenellenbogen, Molecular sensors of estrogen receptor conformations and dynamics, *Mol. Endocrinol.*, 2003, **17**, 2593-2602.

[35] S. Y. Dai, M. J. Chalmers and J. Bruning, Prediction of the tissue-specificity of selective estrogen receptor modulators by using a single biochemical method, *Proc. Natl. Acad. Sci. USA*, 2008, 105, 7171–7176.

[36] L. J. Boerma, G. Xia, C. Qui, B. D. Cox, M. J. Chalmers, C. D. Smith, S. Lobo-Ruppert, P. R. Griffin, D. D. Muccio and M. B. Renfrow, Defining the communication between agonist and coactivator binding in the retinoid X receptor α ligand binding domain, *J. Biol. Chem.*, 2014, **2**, 814-826.

[37] S. Y. Dai, T. P. Burris, J. A. Dodge, C. Montrose-Rafizadeh, Y. Wang, B. D. Pascal, M. J. Chalmers and P. R. Griffin, Unique ligand binding patterns between estrogen receptor alpha and beta revealed by hydrogen-deuterium exchange, *Biochemistry*, 2009, **40**, 9668-9676.

[38] A. Tamrazi, K. E. Carlson, J. R. Daniels, K. M. Hurth and J. A. Katzenellenbogen, Estrogen receptor dimerization: ligand binding regulates dimer affinity and dimer dissociation rate, *Mol. Endocrinol.*, 2002, **12**, 2706-2719.

[39] Y. Arao, K. J. Hamilton, L. A. Coons and K. S. Korach, Estrogen receptor α L543A,L544A mutation changes antagonists to agonists, correlating with the ligand binding domain dimerization associated with DNA binding activity, *J. Biol. Chem.*, 2013, **288**, 21105-21116.

[40] J. Osz, Y. Brélivet, C. Peluso-Iltis, V. Cura, S. Eiler, M. Ruff, W. Bourguet, N. Rochel and D. Moras, Structural basis for a molecular allosteric control mechanism of cofactor binding to nuclear receptors, *Proc. Natl. Acad. Sci. U S A*, 2012, **109**, 588-594.

[41] Y. Arao, K. J. Hamilton, M. K. Ray, G. Scott, Y. Mishina and K. S. Korach, Estrogen receptor AF-2 mutation results in antagonist reversal and reveals tissue selective function of estrogen receptor modulators, *Proc. Natl. Acad. Sci. U.S.A*, 2011, **108**, 14986–14991.

[42] Y. Arao, K. J. Hamilton, E. H. Goulding, K. S. Janardhan, E. M. Eddy and K. S. Korach, Transactivating function (AF) 2-mediated AF-1 activity of estrogen receptor is crucial to maintain male reproductive tractfunction, *Proc. Natl. Acad. Sci. U.S.A*, 2012, **109**, 21140–21145.

[43] S. E. Fawell, J. A. Lees, R. White and M. G. Parker, Characterization and colocalization of steroid binding and dimerization activities in the mouse estrogen receptor, *Cell*, 1990, **60**, 953–962.

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

[45] I. M. Tharun, L. Nieto, C. Haase, M. Scheepstra, M. Balk, S. Möcklinghoff, W. Adriaens, S. A. Dames and L. Brunsveld, Subtype-Specific Modulation of Estrogen Receptor-Coactivator Interaction by Phosphorylation, *Chem. Biol.*, 2014, doi: 10.1021/cb5007097.

[46] D. J. Kojetin and T. P. Burris, Small Molecule Modulation of Nuclear Receptor Conformational Dynamics: Implications for Function and Drug Discovery, *Mol. Pharmacol.*, 2013, **83**, 1–8.

[47] L. Celik, J. D. Lund and B. Schiøtt, Conformational dynamics of the estrogen receptor alpha: molecular dynamics simulations of the influence of binding site structure on protein dynamics, *Biochemistry*, 2007, **46**, 1743-1758.

[48] C. Watanabe, H. Watanabe and S. Tanaka, An interpretation of positional displacement of the helix12 in nuclear receptors: preexistent swing-up motion triggered by ligand binding, *Biochim. Biophys. Acta*, 2010, **1804**, 1832–1840.

[50] D. Kosztin, S. Izrailev and K. Schulten, Unbinding of retinoic acid from its receptor studied by steered molecular dynamics, *Biophys. J.*, 1999, **76**, 188–197.

[51] L. Martínez, I. Polikarpov and M. S. Skaf, Only subtle protein conformational adaptations are required for ligand binding to thyroid hormone receptors: simulations using a novel multipoint steered molecular dynamics approach, *J. Phys. Chem. B*, 2008, **112**, 10741–10751.

[52] W. H. Bisson, R. Abagyan and C. N. Cavasotto, Molecular basis of agonicity and antagonicity in the androgen receptor studied by molecular dynamics simulations, *J. Mol. Graph. Model.*, 2008, **27**, 452–458.

[53] A. Perkins, J. L. Phillips, N. I. Kerkvliet, R. L. Tanguay, G. H. Perdew, S. K. Kolluri and W. H. Bisson, A Structural Switch between Agonist and Antagonist Bound Conformations for a Ligand-Optimized Model of the Human Aryl Hydrocarbon Receptor Ligand Binding Domain, *Biology (Basel).*, 2014, **3**, 645-669.

[54] D. A. Case, T. A. Darden, 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. Goetz, 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, 2012 *AMBER 12*, University of California, San Francisco.

[55] R. Salomon-Ferrer, A. W. Goetz, D. Poole, S. Le Grand and R. C. Walker, Routine microsecond molecular dynamics simulations with AMBER - Part II: Particle Mesh Ewald, *J. Chem. Theory Comput.*, 2013, **9**, 3878-3888.

[56] S. Le Grand, A. W. Goetz and R. C. Walker, SPFP: Speed without compromise - a mixed precision model for GPU accelerated molecular dynamics simulations, *Comp. Phys. Comm.*, 2013, **184**, 374-380.

[57] J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.*, 1977, **23**, 327–341.

[58] H. G. Petersen, Accuracy and efficiency of the particle-mesh-ewald method, *J. Chem. Phys.*, 1995, **103**, 3668–3679.

[59] J. Wereszczynski and J. A. McCammon, Nucleotide-dependent mechanism of Get3 as elucidated from free energy calculations, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 7759-7764.

[60] D. Hamelberg, J. Mongan and J. A. McCammon, Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules, *J. Chem. Phys.*, 2004, **120**, 11919-11929.

[61] D. Hamelberg, C. A. F. de Oliveira and J. A. McCammon, Sampling of slow diffusive conformational transitions with accelerated molecular dynamics, *J. Chem. Phys.*, 2007, **127**, 155102-155109.

[62] B. R. Miller III, T. D. McGee, J. M. Swails, N. Homeyer, H. Gohlke and A. E. Roitberg, MMPBSA. py: An efficient program for end-state free energy calculations, *J. Chem. Theory Comput.*, 2012, **8**, 3314–3321.

[63] I. Kouskoumvekaki, R. K. Petersen, F. Fratev, O. Taboureau, T. E. Nielsen, T. I. Oprea, S. B. Sonne, E. N. Flindt, S. Ó. Jónsdóttir and K. Kristiansen, Discovery of a novel selective PPARγ ligand with partial agonist binding properties by integrated in silico/in vitro work flow, *J. Chem. Inf. Model.*, 2013, **53**, 923-937.

[64] F. Fratev, S. Ó. Jónsdóttir and I. Pajeva, Structural insight into the UNC-45-myosin complex, *Proteins*, 2013, **81**, 1212–1221.

[65] F. Fratev and S. Ó. Jónsdóttir, The phosphorylation specificity of B-RAF(WT), B-RAF(D594V), B-RAF(V600E) and B-RAF(K601E) kinases: An in silico study, *J. Mol. Graph. Model.*, 2010, **28**, 598-603.

[66] F. Fratev, E. Lo Piparo, E. Benfenati and E. Mihaylova, Toxicity study of allelochemical-like pesticides by a combination of 3D-QSAR, docking, Local Binding Energy (LBE) and GRID approaches, *SAR QSAR Environ. Res.*, 2007, **18**, 675-692.

[67] F. Fratev, E. Mihaylova and I. Pajeva, Combination of Genetic Screening and Molecular Dynamics as a Useful Tool for Identification of Disease-Related Mutations: ZASP PDZ Domain G54S Mutation Case, *J. Chem. Inf. Model.*, 2014, **54**, 1524-1536.

[68] D. R. Roe and T. E. Cheatham III, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.

[69] W. R. Adrian, S. Shakhnovich and E. I. Shakhnovich, Contribution of Selection for Protein Folding Stability in Shaping the Patterns of Polymorphisms in Coding Regions, *Mol. Biol. Evol.*, 2014, **31**, 165–176.

[70] Y. Miao, W. Sinko, L. Pierce, D. Bucher, R. C. Walker and J. A. McCammon, Improved reweighting of accelerated molecular dynamics simulations for free energy calculation, *J. Chem. Theory Comput.*, 2014, **10**, 2677-2689.

[71] W. Sinko, C. A. de Oliveira, L. C. Pierce and J. A. McCammon, Protecting High Energy Barriers: A New Equation to Regulate Boost Energy in Accelerated Molecular Dynamics Simulations, *J. Chem. Theory Comput.*, 2012, **8**, 17-23.

[72] T. Shen and D. Hamelberg, A statistical analysis of the precision of reweighting-based simulations, *J. Chem. Phys.*, 2008, **129**, 034103.

[73] P. R. Markwick and J. A. McCammon, Studying functional dynamics in bio-molecules using accelerated molecular dynamics, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20053-20065.

[74] R. Galindo-Murillo, D. R. Roe and T. E. Cheatham III, Convergence and reproducibility in molecular dynamics simulations of the DNA duplex d (GCACGAACGAACGAACGC), *Biochim. Biophys. Acta*, 2015, **1850**, 1041-1058.

[75] C. Bergonzo, N. M. Henriksen, D. R. Roe, J. M. Swails, A. E. Roitberg and T. E. Cheatham III, Multidimensional Replica Exchange Molecular Dynamics Yields a Converged Ensemble of an RNA Tetranucleotide, *J. Chem. Theory Comput.*, 2014, **10**, 492-499.

[76] D. R. Roe, C. Bergonzo and T. E. Cheatham III, Evaluation of enhanced sampling provided by accelerated molecular dynamics with Hamiltonian replica exchange methods, *J. Phys. Chem. B.*, 2014, **118**, 3543-3552.

[77] J. Zhang, M. Qin and W. Wang, Folding mechanism of beta-hairpins studied by replica exchange molecular simulations, *Proteins*, 2006, **62**, 672-685.

[78] W. Wang, W. X. Xu, Y. Levy, E. Trizac and P. G. Wolynes, Confinement effects on the kinetics and thermodynamics of protein dimerization, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 5517-5522.

[79] W. Huang, K. M. Ravikumar and S. Yang, A Newfound Cancer-Activating Mutation Reshapes the Energy Landscape of Estrogen-Binding Domain, *J. Chem. Theory Comput.*, 2014, **10**, 2897–2900.

[80] K. W. Nettles, J. B. Bruning, G. Gil, E. E. O'Neill, J. Nowak, Y. Guo, Y. Kim, E. R. DeSombre, R. Dilis, R. N. Hanson, A. Joachimiak and G. L. Greene, Structural plasticity in the oestrogen receptor ligand-binding domain, *EMBO Rep.*, 2007, **8**, 563-568.

[81] D. Jereva, F. Fratev, I. Tsakovska, P. Alov, T. Pencheva and I. Pajeva, Molecular Dynamics Simulation of the Human Estrogen Receptor Alpha: Contribution to the Pharmacophore of the Agonists, *Biomath Communications*, 2014, **1**, 23-27.

[82] D. G. Teotico, M. L. Frazier, F. Ding, N. V. Dokholyan, B. R. Temple and M. R. Redinbo, Active nuclear receptors exhibit highly correlated AF-2 domain motions, *PLoS Comput. Biol.*, 2008, **4**, 1000111.

[83] S. Movérare-Skrtic, A. E. Börjesson, H. H. Farman, K. Sjögren, S. H. Windahl, M. K. Lagerquist, A. Andersson, A. Stubelius, H. Carlsten, J. Å. Gustafsson and C. Ohlsson, The estrogen receptor antagonist ICI 182,780 can act both as an agonist and an inverse agonist when estrogen receptor α AF-2 is modified, *Proc. Natl. Acad. Sci. U S A*, 2014, **111**, 1180-1185.

[84] M. T. Sonoda, L. Martínez, P. Webb, M. S. Skaf and I. Polikarpov, Ligand dissociation from estrogen receptor is mediated by receptor dimerization: evidence from molecular dynamics simulations, *Mol. Endocrinol.*, 2008, **7**, 1565-1578.

[85] S. Chakraborty, A. S. Levenson and P. K. Biswas, Structural insights into Resveratrol's antagonist and partial agonist actions on estrogen receptor alpha, *BMC. Struct. Biol.*, 2013, **13**, 27.

[86] S. Chakraborty and P. K. Biswas, Structural insights into selective agonist actions of tamoxifen on human estrogen receptor alpha, *J. Mol. Model.*, 2014, **20**, 2338.

[87] K. E. Carlson, I. Choi, A. Gee, B. S. Katzenellenbogen and J. A. Katzenellenbogen, Altered ligand binding properties and enhanced stability of a constitutively active estrogen receptor: evidence that an open pocket conformation is required for ligand interaction, *Biochemistry*, 1997, **36**, 14897–14905.

[88] R. L. Rich, L. R. Hoth, K. F. Geoghegan, T. A. Brown, P. K. LeMotte, S. P. Simons, P. Hensley and D. G. Myszka, Kinetic analysis of estrogen receptor/ligand interactions, *Proc. Natl. Acad. Sci. U S A*, 2002, **99**, 8562-8567.

Movies

Movie 1: H12 conformational sampling in an ERα monomer based on simulation aMDm2 which describes the observed and presented in Figure 1 activation helix populations under and above H11, respectively.

Movie 2: Transition of H12 from an unfolded form to an antagonist and then an antagonist conformation via Path 1

Movie 3: Transition of H12 from an unfolded form to an antagonist via Path 2

Movie 4: Transition of H12 from an unfolded form to an antagonist conformation via Path 1