Cross-talk between the ligand- and DNA-binding domains of estrogen receptor

ARTICLE *in* PROTEINS STRUCTURE FUNCTION AND BIOINFORMATICS · NOVEMBER 2013

Impact Factor: 2.63 · DOI: 10.1002/prot.24331 · Source: PubMed

4 AUTHORS, INCLUDING:

Wei Huang
Case Western Reserve University
16 PUBLICATIONS **98** CITATIONS
SEE PROFILE

Geoffrey L Greene
University of Chicago
205 PUBLICATIONS **18,460** CITATIONS
SEE PROFILE

Sichun Yang
Case Western Reserve University
24 PUBLICATIONS **674** CITATIONS
SEE PROFILE

PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Cross-talk between the ligand- and DNA-binding domains of estrogen receptor

Wei Huang,[1] Geoffrey L. Greene,[2] Krishnakumar M. Ravikumar,[1] and Sichun Yang [1]*

[1] Center for Proteomics and Department of Pharmacology, Case Western Reserve University, Cleveland, Ohio 44106-4988

[2] Ben May Department for Cancer Research, University of Chicago, Chicago, Illinois 60637

ABSTRACT

Estrogen receptor alpha (ERα) is a hormone-responsive transcription factor that contains several discrete functional domains, including a ligand-binding domain (LBD) and a DNA-binding domain (DBD). Despite a wealth of knowledge about the behaviors of individual domains, the molecular mechanisms of cross-talk between LBD and DBD during signal transduction from hormone to DNA-binding of ERα remain elusive. Here, we apply a multiscale approach combining coarse-grained (CG) and atomistically detailed simulations to characterize this cross-talk mechanism via an investigation of the ERα conformational landscape. First, a CG model of ERα is built based on crystal structures of individual LBDs and DBDs, with more emphasis on their interdomain interactions. Second, molecular dynamics simulations are implemented and enhanced sampling is achieved via the "push-pull-release" strategy in the search for different LBD–DBD orientations. Third, multiple energetically stable ERα conformations are identified on the landscape. A key finding is that estradiol-bound LBDs utilize the well-described activation helix H12 to pack and stabilize LBD–DBD interactions. Our results suggest that the estradiol-bound LBDs can serve as a scaffold to position and stabilize the DBD–DNA complex, consistent with experimental observations of enhanced DNA binding with the LBD. Final assessment using atomic-level simulations shows that these CG-predicted models are significantly stable within a 15-ns simulation window and that specific pairs of lysine residues in close proximity at the domain interfaces could serve as candidate sites for chemical cross-linking studies. Together, these simulation results provide a molecular view of the role of ERα domain interactions in response to hormone binding.

Proteins 2013; 81:1900–1909.

Key words: estrogen receptor; protein–protein interaction; CG model; energy landscape; MD.

INTRODUCTION

Estrogen receptor alpha (ERα), a member of the nuclear hormone receptor (NHR) family, plays a key role in eukaryotic transcriptional regulation.[1–4] Once activated by hormones such as estradiol, ERα can bind specific DNA sequences as a homodimer and modulate the expression of genes involved in the development, homeostasis, and metabolism.[5–9] Because ERα is a key target for therapeutic intervention,[10–12] it is vital to obtain a better molecular understanding of ERα action that results from hormone signaling.

ERα is a multidomain protein fold containing a DNA-binding domain (DBD) and a ligand-binding domain (LBD).[2,13] Their individual crystal structures have been determined (as illustrated in Fig. 1). These two domains are connected by a 50-residue flexible hinge region that allows these domains to move freely in three-dimensions (3D). It has been observed that the DBD alone can bind a double-strand DNA,[15–17] and that this DNA binding of the receptor is tighter in the presence of LBD when bound to estradiol.[6] At a structural level, this estradiol binding is also coupled with a major rearrangement of the critical activation helix H12 at the C-terminus, which is positioned on the LBD surface and covers the estradiol binding pocket.[18–22]

The mechanisms by which ERα DBDs and LBDs interact with each other when ERα is bound to DNA are still unclear. For several other members of the NHR family, it has been demonstrated that the LBD–DBD interface plays an allosteric role to mediate the signal transduction between ligand binding and DNA association.[23–25] For example, a single point mutation at the interface of the

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Cleveland Foundation and the Department of Defense Breast Cancer Research Program; grant number: W81XWH-11-1033.

*Correspondence to: Sichun Yang, Ph.D., Case Western Reserve University 10900 Euclid Ave., BRB 929 Cleveland, OH 44106-4988. Email: sichun.yang@case.edu

Received 9 January 2013; Revised 22 April 2013; Accepted 9 May 2013
Published online 5 June 2013 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24331

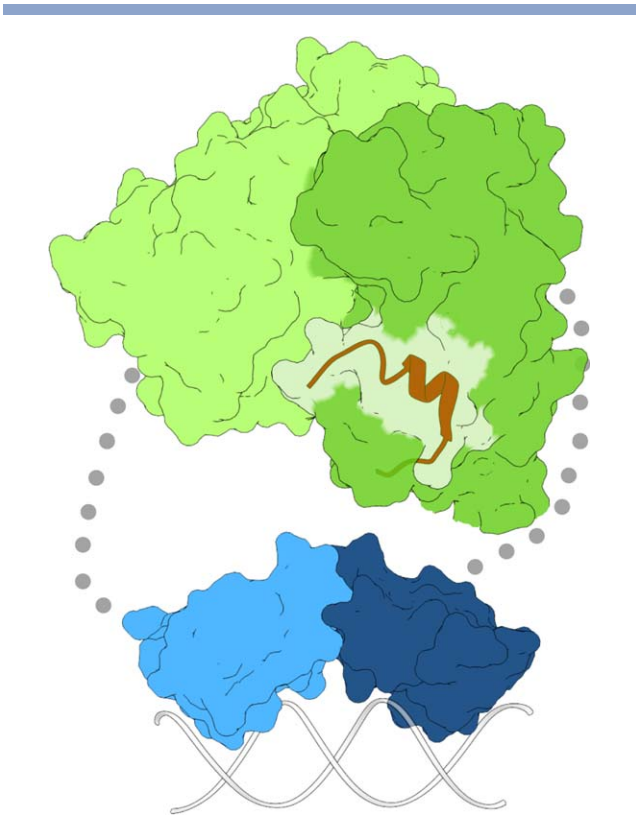

Figure 1

Schematic diagram of ERα functional domains of a LBD and a DBD. LBD and DBD dimers are displayed in green and blue, respectively, and their connecting hinge loops are marked by gray dots. The DBD-specific DNA duplex is shown as ribbon in white. The activation helix H12 is highlighted in red. Part of this figure was rendered using UCSF Chimera.[14] [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

peroxisome proliferator-activated receptor (PPAR)/retinoid X receptor (RXR) complex affects its DNA binding properties to a PPAR response element.[21] More recently, mutagenesis studies in the androgen receptor (AR) show that several mutations in the AR LBD can decrease DNA binding affinity, and conversely, mutations in the DBD also affect ligand binding.[25] These observations strongly suggest the internal communication between LBDs and DBDs is required to facilitate *in vitro* DNA binding. For ERα, how these domains interact and communicate remains largely unknown, in part due to the lack of structural information of a full LBD–DBD/DNA complex and the lack of knowledge about the LBD–DBD interactions at a molecular level.

In this study, we used a computational approach of coarse-grained (CG) and subsequent all-atom molecular dynamics (MD) simulations to explore the physical interactions between the ERα DBD and LBD domains from an energy landscape perspective. Specifically, a CG model and an efficient sampling method were implemented to simulate the ERα conformational landscape. The projection of simulation data onto a 3D globe allows one to picture the energy landscape of ERα multidomain interactions. Multiple ERα conformational states were identified from the landscape, highlighting different modes of LBD–DBD interactions. Final structural refinement via all-atom simulations starting from CG-predicted models provides detailed structural features at the domain interfaces that could be experimentally tested. Overall, these results provide new speculation into ERα domain cross-talk mechanisms responsible for the allosteric control of receptor binding to DNA.

METHODS

Details of the CG model

A CG model was implemented where each amino acid of ERα is represented by a single bead positioned at its Cα atom and each nucleotide of DNA is represented by its O5′ atom. The CG energy function for ERα (E_{181}-P_{552}) is defined for each component as follows. The energy functions for the estradiol-bound LBD dimer (S_{309}-P_{552}) and the DBD dimer (E_{181}-K_{252}, in complex with a 17-bp double-stranded DNA) were built based on their crystal structures (protein data bank (PDB) entries 1QKU and 1HCQ),[16,21] using the widely used Gō-type potential. The hinge region connecting DBD and LBD (G_{253}-L_{308}) was built using loop modeling method in MODELLER.[26] Bonded interactions, including bond, angle, and dihedral terms, were modeled using typical energy functions as done previously (see details in Refs. [27,28]. Nonbonded attractive interactions between native contact pairs between LBDs and DBDs were modeled using Lennard-Jones (LJ) interactions, $E_{\mathrm{LJ}} = \sum_{i,j} \varepsilon_o [5(\sigma_{ij}^o / r_{ij})^{12} - 6(\sigma_{ij}^o / r_{ij})^{10}]$, where r_{ij} is the residue–residue distance and σ_{ij}^o is the corresponding distance in the reference structure. The contact pairs within each domain were defined using the CSU software.[29] Native-like DBD–DNA interactions were modeled using $\varepsilon_0 = 5$ kcal/mol whenever residue-nucleotide pairs were within a distance of 15 Å.

The interactions between LBDs and DBDs are specified by the energy function $E_{\mathrm{LBD\text{-}DBD}}$, which includes electrostatic (E_{elec}) and hydrophobic (E_{H}) components,

$$E_{\mathrm{LBD-DBD}} = E_{\mathrm{elec}} + E_{\mathrm{H}}. \quad (1)$$

Here, $E_{\mathrm{elec}} = \sum_{i,j} q_i q_j / (4\pi\varepsilon_o D_{\mathrm{eff}} r_{ij})$ was used where q_i is the charge of residue i and ε_0 is the vacuum electric permittivity. An effective dielectric coefficient $D_{\mathrm{eff}} = D_s \exp(r_{ij}/\xi)$ is applied to reflect the shielding effect between two residues separated by a distance of r_{ij}. $D_s = 10$ was used to describe the local dielectric environment when two domains are forming a close interface and $\xi = 8.2$ Å to mimic the screening effect at a 150-mM salt concentration. At pH 7, residue charges $q_i = +e$ for Lys and Arg, $-e$ for Asp and Glu, and $+0.5\,e$ for His (e is the elementary charge) were used.[27,30] Hydrophobic

interactions (E_H) are either attractive (LJ-type) ($\varepsilon_{ij} < 0$) or purely repulsive ($\varepsilon_{ij} \geq 0$) where $\varepsilon_{ij}=\alpha(e_{ij}^{MJ}+\beta)$ between residues i, j and e_{ij}^{MJ} is the Miyazawa-Jernigan statistical energy.[31] The values of $\alpha = 0.4$ and $\beta = 1.3$ were used to balance attractive and repulsive interactions.[27] We used $E_H(i,j)=|\varepsilon_{ij}|[5(\sigma_{ij}/r_{ij})^{12}-6(\sigma_{ij}/r_{ij})^{10}]$ if $\varepsilon_{ij} < 0$, and $E_H(i,j)=\varepsilon_{ij}[5(\sigma_{ij}/r_{ij})^{12}(1-\exp(-(r_{ij}-\sigma_{ij})/d)^2)]$ if $\varepsilon_{ij} \geq 0$, where $d = 3.8$ Å. The distance σ_{ij} is defined by

$$\sigma_{ij}=\gamma(r_i+r_j) \quad (2)$$

where the value of $\gamma = 0.625$ was chosen[27] and r_i is the van der Waals radius of residue i.[30] The non-native interactions of DNA with LBDs and the hinge region were mostly repulsive with $\varepsilon_{ij} = 1$ kcal/mol.

Details of the "push-pull-release" (PPR) sampling strategy

The CG model was implemented via Langevin MD simulations in a modified version of the CHARMM program.[32] Simulations were performed at 300 K with a friction coefficient of 50 ps^{-1}.[33] A simulation time step of 10 fs was used and simulated coordinates were saved every 100 ps. A sampling strategy of PPR was implemented via a biasing potential E_{PPR},[27]

$$E_{PPR}=\begin{cases} 0, & \text{when } R_t \leq R_c \text{ and } r_{min} < r_o, \\ k(R-R_t)^2, & \text{otherwise.} \end{cases} \quad (3)$$

where R is the instantaneous center-of-mass distance between two selected domains, R_t is its corresponding target distance, and r_{min} is the closest residue–residue distance between the two domains. The two domains move between $R_{min} \leq R_t \leq R_{max}$, where the values of $R_{min} = 0$ Å and $R_{max} = 100$ Å were used. This E_{PPR} potential is applied to each PPR cycle with a simulation time of L and repeated for N times, where each cycle includes three steps: (1) "push" the two domains away from each other when they are close, (2) "pull" the domains closer until they are separated by a threshold distance R_c, and (3) "release" them to interact freely by removing the biasing potential. For ERα simulations, the parameters $N = 10$, $L = 25$ ns, $k = 100$ kcal/(mol·Å^2), $r_o = 7.6$ Å, and $R_c = 45$ Å were used. Four sets of PPR-assisted simulations were performed, each accounting for a different biasing pair between two LBDs and two DBDs. Within each set, a total of 145 simulation runs were carried out simultaneously, each with a different initial configuration of LBD–DBD relative orientation.[27] A simulation time of 145 μs was achieved in total.

Details of orientation-based clustering analysis

A structural clustering analysis was applied mainly based on relative orientations between domains. First, structural alignment was performed on LBDs only. Second, resultant configurations of DBDs were used to calculate pair-wise root-mean-square deviation (RMSD) values. Finally, these RMSD values were used as an input for a hierarchical clustering analysis as implemented in MATLAB. This clustering analysis was applied to selected simulation snapshots if they have the lowest $E_{LBD\text{-}DBD}$ from the last 5-ns segment in each PPR cycle, where $E_{LBD\text{-}DBD} < 0$ kcal/mol, and the RMSD values of individual DBD and LBD monomers are within 1.5 and 3 Å to their corresponding crystal structures, respectively. In addition, mirror conformations by exchanging the coordinates of two protein chains were generated due to ERα homodimeric symmetry. Thus, a total number of 9780 configurations were used for the clustering analysis.

Details of all-atom simulations

All-atom MD simulations were performed starting from CG-predicted conformations. Initial ERα models were built based on crystal structures of individual domains. First, the CG models were used as the templates to build atomically detailed structures, where crystal structures of the LBDs (PDB entry 1QKU[21]) and DBDs (PDB entry 1HCQ[16]) were aligned on the top of Cα atoms of CG models. Second, the restraints of Cα atoms within domains toward CG models were achieved via targeted MD simulations.[34] An implicit solvent GB model was used to reduce the computational cost[35] and a total of 40-ns simulation time was performed to reach a Cα-RMSD value of 0.5 Å (relative to the CG models of ERα domains). Finally, standard all-atom simulations without any bias were performed for model refinement using the program NAMD,[36] where the force fields of AMBER99SB[37] and AMBER99bsc0[38,39] were used. Parameters for the estradiol were generated using GAFF[40] and ANTECHAMBER.[41–43] All parameters were prepared using the AMBER LEaP module.[44] The system was placed in a rectangle water box with more than 12 Å padding and 16 charge-neutralizing sodium ions with additional salts (Na+, Cl−) around the solute resulting in 150 mM salt condition. Langevin simulations with a total of 15 ns were performed for each state at a pressure of 1 atm and a temperature of 300 K.

RESULTS AND DISCUSSION

A cartoon representation of the ERα multidomain architecture is shown in Figure 1, illustrating ERα domains for both LBD and DBD. Here, a recently developed CG MD simulation pipeline was applied to simulate ERα dynamics.[27] Specifically, the search for different LBD–DBD assemblies was enhanced via a sampling strategy of PPR. After grouping similar configurations into a conformational state via a hierarchical clustering scheme, ERα simulation data were projected onto a 3D globe

specifying the ERα interdomain orientation, and the interacting energy of each state was used to visualize the LBD–DBD energy landscape and identify favorable conformations accessible on the landscape. Finally, the structural features of stable ERα conformational states were illustrated to highlight their interdomain interactions.

CG simulations with the PPR sampling strategy

The CG model implemented in ERα simulations has two main components: intradomain and interdomain interactions. For each component of the LBD–DBD complex, a Gō-like model was used to model the dynamics within both the LBD dimer and the DBD dimer by taking advantage of the availability of their crystal structures.[16,21] This treatment allowed each dimer to structurally fluctuate around its native conformation, as well as maintain substantive flexibility as required for conformational deformation in an encounter complex. For the interdomain LBD–DBD interactions, a residue-specific energy function of $E_{\text{LBD-DBD}}$ [Eq. (1)] was implemented by accounting for explicit electrostatic and hydrophobic interactions (see details in Methods Section). Thus, a CG model of ERα was built based on crystal structures as well as physical principles for the purpose of MD simulations.

The search for ERα conformations was enhanced via the sampling strategy of PPR. This PPR approach was adopted to facilitate the dissociation and reassociation between LBDs and DBDs (see Ref. [27] and Methods Section). This sampling enhancement was achieved by applying a biasing potential to the center-of-mass distance of the two domains [Eq. (2)], which can push the domains together to associate, pull them away to disassociate, and release them to relax by removing the biasing potential when the domains are encountered. It should be noted that during the pulling and pushing, two domains were allowed to rotate freely, thus helping achieve broader sampling in their relative orientations.

To illustrate the sampling coverage achieved by PPR, a 3D sphere was used to specify LBD–DBD orientations. The LBDs were first used for the alignment of each simulation snapshot and placed in the center of this globe. The center-of-mass of DBDs was then projected onto the surface of the globe [Fig. 2(A)]. Compared to the simulations without using PPR, Figure 2(B) shows that within a same simulation window of 125 ns, the sampling efficiency was considerably enhanced by PPR with a much larger surface area spanned on the globe. The sampling of these PPR-assisted simulations was further accelerated by launching a set of 145 parallel CGMD runs each with a different initial configuration, as illustrated in Supporting Information Figure S1(A). In addition, four sets of such simulations were performed, each with a different biasing potential between a pair of monomeric LBD and DBD (see Methods Section). As a result, a total of 145-μs simulation time was achieved and an accumulated 61.625 μs from unbiased PPR segments was used for the rest of the analyses.

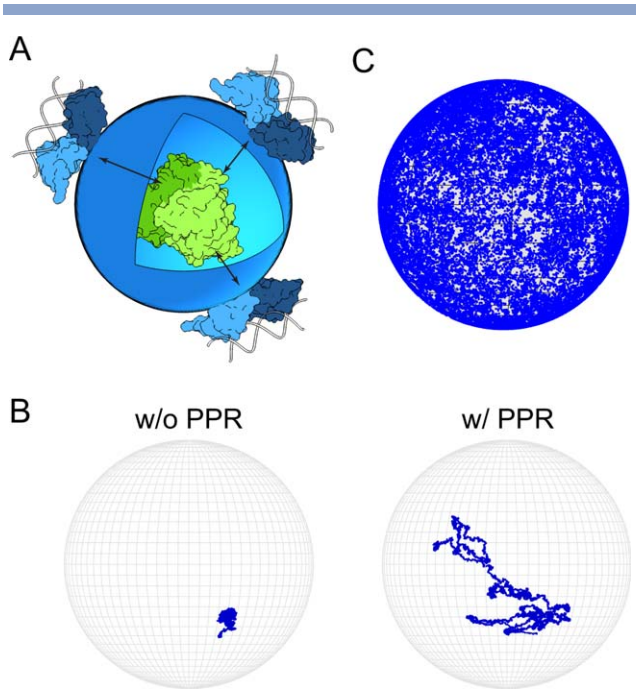

Figure 2

Sampling coverage using PPR in CGMD simulations. (**A**) Illustration of the 3D globe specifying LBD–DBD orientations, where LBDs are in the center (in green) and DBDs are on the surface (in blue). (**B**) Sampling coverage without and with using PPR. The globe is projected with a single simulation trajectory without (left) and with (right) the assistance of PPR. In both simulations, a same initial configuration and a simulation length of 125 ns were used. (**C**) Total sampling coverage from all trajectories projected on the globe where each simulation snapshot is represented by a blue dot. Front view is shown and alternative views are also available in Supporting Information Figure S1. The projected snapshots were taken only from free and released portions of PPR simulation cycles with an energy cutoff of $E_{\text{LBD-DBD}} < 0$ kcal/mol and an ensemble of 10 structures from each cycle. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 2(C) illustrates the sampled area of the globe covered by simulation snapshots, where each snapshot is represented by a blue dot. Clearly, most of the surface area was sampled by PPR-assisted simulation trajectories, suggesting that sufficient sampling of various LBD–DBD orientations was obtained. We note that both front and back views of the globe were based on the placement of LBDs in the center [Supporting Information Fig. S1(B)]; an alternative placement of DBDs in the center is also shown in Supporting Information Figure S1(C), where more than half of the sphere is well covered (and the other half is not covered due to the blockage of its DNA binding). Together, these results suggest that this PPR strategy substantially assisted the CGMD simulations with a rather comprehensive sampling coverage in the ERα domain orientational space.

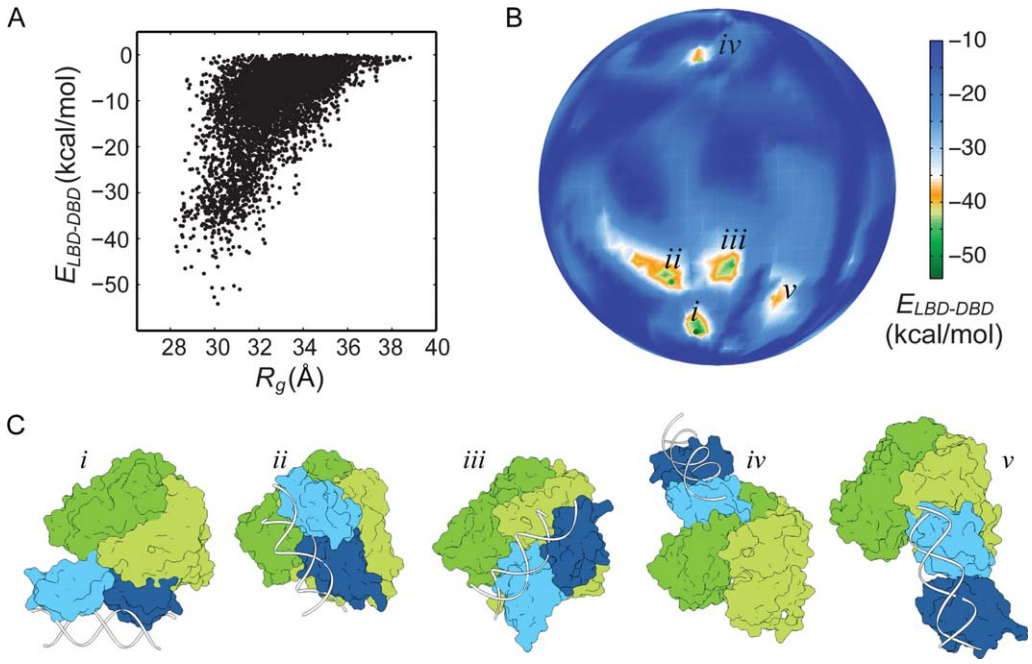

Figure 3

Energy landscape of ERα multidomain interactions. (**A**) A plot of LBD–DBD interacting energy versus ERα-DNA radius of gyration R_g. A total of 4890 snapshots were used for the plot. (**B**) An energy globe for the projection of simulation data and the identification of energetically stable ERα conformations. The globe is colored based on the energies of $E_{LBD-DBD}$, as indicated by the color bar on the right. From this energy plot, there are five conformational states *i*–*v* that are identified as energetically stable and ranked in the order of their $E_{LBD-DBD}$ values. Their $E_{LBD-DBD}$ and R_g values are −51.6 ± 1.3 kcal/mol and 30.40 ± 0.17 Å for state *i*, −44.2 ± 1.6 kcal/mol and 29.63 ± 0.10 Å for state *ii*, −44.4 ± 1.1 kcal/mol and 28.60 ± 0.08 Å for state *iii*, −39.8 ± 1.9 kcal/mol and 31.58 ± 0.07 Å for state *iv*, and −37.2 ± 1.3 kcal/mol and 31.72 ± 0.33 Å for state *v*. A total of 9780 snapshots were used for the projection, which is doubled due to the homodimeric symmetry. (**C**) Surface view of a representative structure in each ERα conformational state. For clear visualization, their hinge loops are not shown in these surface views. An ensemble of structures for each state is also shown in Supporting Information Figure S4.

Energy landscape on a globe: clustering and projection

The visualization of ERα conformational landscape was analyzed by projecting simulation data onto a 3D globe after clustering.[27] First, a hierarchical clustering algorithm was implemented to organize the large amount of simulation data, which was based on pair-wise RMSDs of the DBDs after LBDs were aligned (see Methods Section). We also note that only the last 5-ns segment within the "release" part of each PPR cycle was used for the analysis because a period of equilibration time before this segment was observed to be helpful for the relaxation of the configurations coming out of PPR pushing segments, as shown in Supporting Information Figure S2. As a result, a total of 4890 structures were selected from the entire CGMD simulations, each representing the lowest $E_{LBD-DBD}$ configuration from its corresponding PPR cycle. Figure 3(A) shows a plot of $E_{LBD-DBD}$ versus radius of gyration (R_g) for these configurations. Clearly, a broad range of configurations was observed in the simulations, ranging in $E_{LBD-DBD}$ from 0 to –54 kcal/mol and R_g from 28 to 39 Å.

Second, the energetics of ERα conformational landscape was achieved by the projection of resulting clusters onto an energy globe. Specifically, this globe was colored based on the lowest $E_{LBD-DBD}$ value within each cluster. Figure 3(B) illustrates the projected energy globe, characterizing the ERα overall energy landscape. This characterization was achieved with a clustering analysis procedure with a total number of clusters $N = 300$. Additional analyses using $N = 500$ and 1000 yielded a similar overall picture of the energy-colored globe [Supporting Information Fig. S3(A)]. Thus, the globe resulting from $N = 300$ clusters was used to complete the analysis. The lowest-energy configurations in the top 25 and 30% of all unbiased simulations were also projected onto the energy globe [Supporting Information Fig. S3(B)]. These results show that each state remains at a similar energy basin when different pools of configurations are used in the projection, suggesting that these CG-identified states are capable of representing and illustrating the simulated landscape. Overall, compared to traditional 2D plots as shown in Figure 3(A), this energy globe analysis offers an advantage for the navigation of a complex

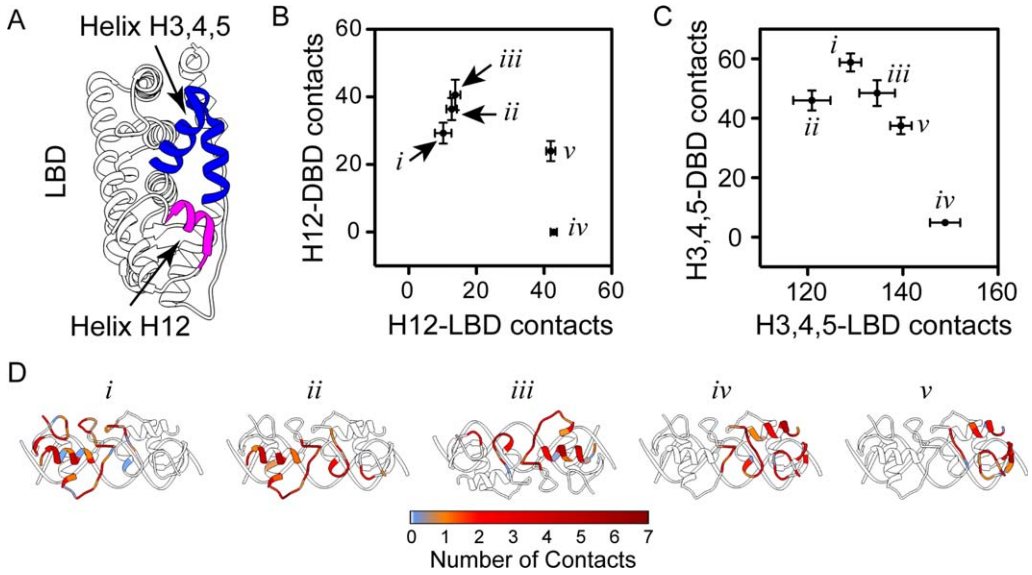

Figure 4

ERα LBD–DBD interfaces. (**A**) Two LBD segments: helices H3–H5 (residue 350–385) and helix H12 (residue 536–544), highlighted in blue and magenta, respectively. (**B**) The interactions of helix H12 with LBDs and DBDs. The calculations of H12-LBD contacts was based on the estradiol-bound LBD crystal structure (PDB entry 1QKU),[20] and the calculations of H12-DBD contacts was based on the criterion that the instantaneous distance is within a factor of 1.2 of the corresponding distance for each pair of amino acids as defined in Eq. (2). An ensemble of 10 highest ranked configurations based on $E_{LBD\text{-}DBD}$ was used to calculate the standard deviation (marked by error bars) for each number of contacts. (**C**) The interactions of helices H3–H5 with LBDs and DBDs. (**D**) DBD surface involved in LBD–DBD interactions. Colored are the DBD surface residues according to the number of contacts for each residue (marked by the color bar below). Results for the LBD surfaces are also shown in Supporting Information Figure S5(A) and S6(A). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

conformational space and for the identification of stable ERα conformations.

Identification of multiple ERα conformations

The energy-landscape-on-a-globe analysis readily reveals that multiple ERα conformational states are physically accessible via CGMD simulations. Figure 3(B) outlines five energetically favorable ERα states (marked by *i*–*v*), which can be easily identified from this colored globe. Clearly, these states differ not only in their LBD–DBD interacting energies, but also in their domain interfaces. Based on their surface views [Fig. 3(C)] and structural ensembles (Supporting Information Fig. S4), both states *iv* and *v* are in a more extended conformation with larger R_g values than states *i*–*iii*. Within these two states, only one monomer, either LBD or DBD, is involved in LBD–DBD interactions, while both DBDs make contact with both LBDs in three other states. Notably, such a large difference in overall size, for example, about 3 Å between state *iii* and state *iv*/*v*, could be experimentally characterized by small-angle X-ray scattering (SAXS).

Among these observed states, different LBD–DBD binding interfaces are formed. For example, state *iv* is located at the upper region of the globe, which has a distinct LBD binding surface compared to the rest of four states [Figs. 3(C) and Supporting Information S5(A)]. It should be noted that this interface is observed to be partially involved in the interactions of ERα LBD with a peptide antagonist [Supporting Information Fig. S5(B)].[45,46] A closer look at the LBD surface shows that a similar region is shared and utilized for the formation of close LBD–DBD interfaces at states *i*–*iii* and *v* [Supporting Information Fig. S6(A)], where mostly one LBD monomer participates in the interfacial interactions. From the LBD side, this interface mainly consists of two distant sequence segments: helices H3–H5 and helix H12, marked in blue and magenta, respectively [Fig. 4(A)]. A portion of these LBD surfaces also form the binding site for the interactions with a coactivator peptide[20] [Supporting Information Fig. S6(B)] as well as for interactions with 4-hydroxytamoxifen observed in ERβ.[47]

One striking feature is that H12 is crucial in mediating LBD–DBD interactions in several conformational states. Figure 4(B) shows that H12 forms a large number of contacts with DBD at states *i*–*iii*, and *v*, and most of them contribute to the predominant hydrophobic interdomain interactions (Supporting Information Fig. S7). This involvement of H12 was further tested by mutating several charged and hydrophobic residues in helix H12 (Supporting Information Fig. S8). The destabilization of domain interactions is clearly reflected by the loss of interacting energies, suggesting that H12 plays a

structural role in gluing and stabilizing interdomain interactions. Furthermore, compared with previous crystallographic studies on LBDs alone, in which this H12 helix is structurally reoriented upon estradiol binding,[18,20,21] the H12 involved in domain interfaces maintains this estradiol-bound orientation in states *i–iii*. In fact, this particular H12 adopts a similar estradiol-bound conformation, in which its native contacts with LBD were sacrificed, throughout all of these low $E_{\text{LBD-DBD}}$ states, suggesting that its interactions with DBD can stabilize receptor LBD–DBD conformations. Given the previous observation that the LBD, in the presence of hormone, promotes or stabilizes tighter DNA binding.[6] Our results strongly suggest that H12 plays a key role in receptor stabilization associated with tight DNA binding. This observation is also consistent with the fact that deletion of H12 can eliminate the ability of ERα to bind DNA.[6] Together, these findings indicate that H12 is a key structural modulator in hormone-activated ERα allostery; it mediates internal communication by bridging LBD and DBD association, through which the reorientation of H12, triggered by hormone binding, results in tighter DNA binding.

Another LBD surface involved in LBD–DBD interfaces consists of helices H3–H5 [labeled in blue in Fig. 4(A)]. Figure 4(C) shows that these helices make a large number of contacts with DBD in states *i–iii*, and *v*, while almost no contact occurs in state *iv*. More details, including residue locations and total contacts, on these interfacial sites are also illustrated in Supporting Information Figures S5 and S6. Overall, this H3–H5 surface, together with H12, accounts for the stabilization of dominant LBD–DBD conformations.

It should be noted that one exception to this H12-driven mechanism is state *iv*, in which H12 is not involved in the formation of its LBD–DBD interface. Figure 4(B) shows that nearly no contact is formed between H12 and DBDs at state *iv*. More strikingly, this state has a similar domain orientation regarding LBD interface-forming surfaces as observed in a new crystal structure of the nuclear receptor homodimer HNF4α (PDB entry 4IQR)[48] (Supporting Information Fig. S9), where the top regions of LBD surfaces, instead of the region of H12 and H3–H5, are involved in the interface formation in both conformations. While there is a difference in their corresponding DNA segments (an inverted repeat for ERα and a direct repeat for HNF4α), this result highlights that the simulations are capable of capturing the large-scale domain organizations.

The DBD surfaces involved in forming interfaces vary from state to state. Figure 4(D) highlights those interacting residues on DBD surfaces that form multiple contacts with its LBD partner. Among all of these states, the surface itself mainly consists of a loop connecting the first two helices as well as a large portion of the third helix. The key difference, among states *i–iii*, and *v*, is that different sets of DBD residues are utilized for the interfacial packing, resulting in distinct LBD–DBD orientations. This large-scale difference of domain orientations could be examined by small-angle scattering measurements, which was not attempted here but will be reported in future communications. In addition, both DBD monomers interact with LBDs in states *i–iv* [Fig. 4(D)].

This observation that both DBD monomers can interact with LBDs suggests a possible role of LBD in receptor stabilization. It has been suggested that the interactions between two DBD monomers are relatively weak.[49] Here, based on the simulation results, their interactions with LBDs can be rather strong with a gain of energy, $E_{\text{LBD-DBD}} > 30$ kcal/mol (Fig. 3). This stabilization via LBD interactions is also consistent with the observation that the binding of receptor to DNA is tighter in the presence of hormone-bound LBD compared to DBD alone.[6] These results suggest that hormone-bound LBD dimers can serve as a scaffold to position the DBD dimer, resulting in the stabilization of the DBD–DNA complex and an increase in DNA binding.[49] This notion is further supported by the observed increase of DNA binding affinity when two DBD monomers are connected by an antibody or a linker.[49] Thus, it is likely that the presence of LBDs stabilizes receptor binding to DNA, similar to the antibody or the linker, in which the LBD scaffold is essential for the correct positioning of the DBD dimer.

Refinement by all-atom simulations

To examine the stability and detailed features of ERα structures, all-atom MD simulations were performed starting from the CG-predicted states (see details in Methods Section). Figure 5(A) shows that all five states maintain within a RMSD value of 3 Å (relative to their corresponding starting configurations), suggesting that each state is significantly stable within its 15-ns simulation window. This is further illustrated by the superposition of the CG and all-atom models [Fig. 5(B)]. While large displacement (up to 3 Å) can occur at the loops or domain-connecting hinge loops, little deviation is observed at the domain interfaces or the core regions of each domain. Thus, these all-atom simulations support the overall assessment of predicted conformational states by CG simulations.

Close proximity of lysine residues at the DBD–LBD interfaces are observed via detailed analyses of all-atom simulation data. Figure 5(B) shows that several lysine residues are within a distance of 10 Å (e.g., K_{244}-K_{520} at state *i*, K_{244}-K_{362} at state *ii*, K_{231}-K_{303} or K_{231}-K_{362} at state *iii*, and K_{213}-K_{492} at state *iv*). This observation indicates that these lysine residues could serve as candidate sites of chemical cross-linking for experimental validation. In addition, alternative residues, such as tyrosine and cysteine, could be used in cross-linking experiments

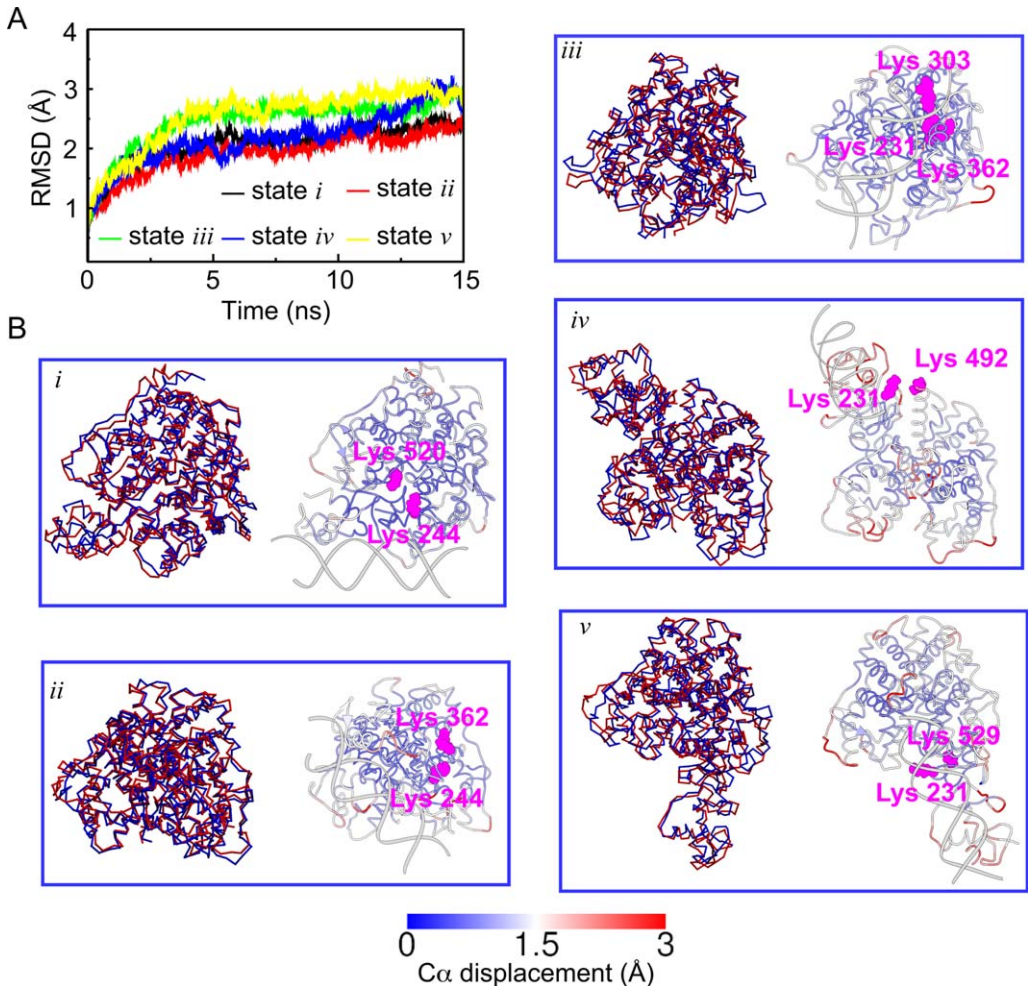

Figure 5

All-atom simulations of five CG-predicted ERα states. (**A**) Simulation trajectories of five conformational states *i*–*v*. RMSDs were calculated based on Cα atoms of the DBD–LBD complex (excluding hinge loop and DNA) with reference to the starting structure [shown in Fig. 3(C)]. (**B**) Comparison of all-atom and CG models. In each panel, a CG model is superimposed with the Cα traces of its all-atom model taken from the last frame of the 15-ns simulation trajectory (left) and the displacement between two models is also colored based on their Cα distances (right). In addition, lysine residues as candidate sites for chemical cross-linking at the domain interfaces are highlighted in magenta. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

based on their pair-wise distances (shown in Supporting Information Fig. S10).

CONCLUSIONS

A generalized PPR-CGMD simulation pipeline has been applied to the investigation of internal interactions between the LBDs and DBDs of ERα. Projection of simulation data onto a 3D energy globe reveals multiple LBD–DBD orientations that are energetically stable, instead of a singular stable conformation. This conformational multiplicity is mainly due to the long LBD–DBD hinge, whose intrinsic flexibility allows each domain to move freely in 3D. Given the increased demands for structural knowledge about multidomain proteins,[50–52] it appears that PPR-CGMD simulations can be broadly applicable to the exploration of domain interactions, for example, within a multidomain complex.

Structural refinement is achieved by using atomic-level simulations starting from CG-predicted structural models. This multiscale approach has provided a stability check for these CG models. It also reveals structural features, especially at the domain interfaces, which could be used for validation designs using SAXS and/or cross-linking measurements.

There are two major findings revealed from these simulations. First, among identified ERα conformations, the activation helix H12 is critical in mediating the cross-talk between LBD and DBD, where four out of five stable states utilize H12 as part of the domain interface. This result suggests that H12 is not only responsive upon

ligand binding, as observed in crystallographic studies of LBD alone,[18,20,21] but also critical for receptor-DNA binding. Second, the stabilization of LBD–DBD interaction via multiple LBD surfaces also suggests a new role of LBD as a scaffold for the stabilization of the DBD–DNA complex, as observed in *in vitro* studies where an increase of DNA binding is observed due to the presence of LBD. Together with previous observations, our findings suggest a molecular model for this critical hormone signaling, where the LBD, when bound to estradiol, stabilizes the DBD dimer for tight DNA binding via LBD surfaces including H12.

Finally, it is worth noting that the focus here was to identify receptor conformations that are energetically stable. This focus was inspired by recent observations that formation of close LBD–DBD interface in other NHR family members,[23,25] as a result of the formation of compact NHR conformations. However, it should also be noted that an elongated/extended receptor conformation has been observed for NHR heterodimers in solution or at cryogenic temperatures.[53,54] The possibility of a similar extended conformation for ERα will be explored in future studies.

ACKNOWLEDGMENT

Computational support was provided by the CWRU High Performance Computing Cluster.

REFERENCES

1. Krust A, Green S, Argos P, Kumar V, Walter P, Bornert J, Chambon P. The chicken oestrogen receptor sequence: homology with v-erba and the human estrogen and glucocorticoid receptors. EMBO J 1986;5:891–897.
2. Kumar V, Green S, Staub A, Chambon P. Localisation of the oestradiol-binding and putative DNA-binding domains of the human oestrogen receptor. EMBO J 1986;5:2231–2236.
3. Kumar V, Green S, Stack G, Berry M, Jin JR, Chambon P. Functional domains of the human estrogen receptor. Cell 1987;51:51–941.
4. Evans RM. The steroid and thyroid hormone receptor superfamily. Science 1988;240:889–895.
5. Notides A, Lerner N, Hamilton D. Positive cooperativity of the estrogen receptor. Proc Natl Acad Sci USA 1981;78:4926–4930.
6. Kumar V, Chambon P. The estrogen receptor binds tightly to its responsive element as a ligand-induced homodimer. Cell 1988;55: 145–156.
7. Wang H, Peters G, Zeng X, Tang M, Ip W, Khan S. Yeast two-hybrid system demonstrates that estrogen receptor dimerization is ligand-dependent in vivo. J Biol Chem 1995;270:23322–23329.
8. Nilsson S, Mkel S, Treuter E, Tujague M, Thomsen J, Andersson G, Enmark E, Pettersson K, Warner M, Gustafsson JA. Mechanisms of estrogen action. Physiol Rev 2001;81:1535–1565.
9. Dahlman-Wright K, Cavailles V, Fuqua SA, Jordan VC, Katzenellenbogen JA, Korach KS, Maggi A, Muramatsu M, Parker MG, Gustafsson JA. International union of pharmacology. LXIV. Estrogen receptors. Pharmacol Rev 2006;58:773–781.
10. Jordan VC. Tamoxifen: a most unlikely pioneering medicine. Nat Rev Drug Discov 2003;2:205–213.
11. Ettinger B, Black DM, Mitlak BH, Knickerbocker RK, Nickelsen T, Genant HK, Christiansen C, Delmas PD, Zanchetta JR, Stakkestad J, Glüer CC, Krueger K, Cohen FJ, Eckert S, Ensrud KE, Avioli LV, Lips P, Cummings SR. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple outcomes of raloxifene evaluation (more) investigators. JAMA 1999;282:637–645.
12. Osborne CK, Zhao H, Fuqua SA. Selective estrogen receptor modulators: structure, #function, and clinical use. J Clin Oncol 2000;18: 3172–3186.
13. Tora L, White J, Brou C, Tasset D, Webster N, Scheer E, Chambon P. The human estrogen receptor has two independent nonacidic transcriptional activation functions. Cell 1989;59:477–487.
14. Pettersen E, Goddard T, Huang C, Couch G, Greenblatt D, Meng E, Ferrin T. UCSF Chimera visualization system for exploratory research and analysis. J Comput Chem 2004;25:1605–1612.
15. Schwabe JW, Chapman L, Finch JT, Rhodes D, Neuhaus D. DNA recognition by the oestrogen receptor: from solution to the crystal. Structure 1993;1:187–204.
16. Schwabe JW, Chapman L, Finch JT, Rhodes D. The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. Cell 1993; 75:567–578.
17. Schwabe JW, Chapman L, Rhodes D. The oestrogen receptor recognizes an imperfectly palindromic response element through an alternative side-chain conformation. Structure 1995;3:201–213.
18. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engstrm O, Ohman L, Greene GL, Gustafsson JA, Carlquist M. Molecular basis of agonism and antagonism in the oestrogen receptor. Nature 1997;389:753–758.
19. Tanenbaum DM, Wang Y, Williams SP, Sigler PB. Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proc Natl Acad Sci USA 1998;95:5998–6003.
20. Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. Cell 1998;95:927–937.
21. Gangloff M, Ruff M, Eiler S, Duclaud S, Wurtz JM, Moras D. Crystal structure of a mutant hERα ligand-binding domain reveals key structural features for the mechanism of partial agonism. J Biol Chem 2001;276:15059–15065.
22. Shiau AK, Barstad D, Radek JT, Meyers MJ, Nettles KW, Katzenellenbogen BS, Katzenellenbogen JA, Agard DA, Greene GL. Structural characterization of a subtype-selective ligand reveals a novel mode of estrogen receptor antagonism. Nat Struct Biol 2002; 9:359–364.
23. Chandra V, Huang P, Hamuro Y, Raghuram S, Wang Y, Burris TP, Rastinejad F. Structure of the intact PPARγ-RXRα nuclear receptor complex on DNA. Nature 2008;456:350–356.
24. Zhang J, Chalmers MJ, Stayrook KR, Burris LL, Wang Y, Busby SA, Pascal BD, Garcia-Ordonez RD, Bruning JB, Istrate MA, Kojetin DJ, Dodge JA, Burris TP, Griffin PR. DNA binding alters coactivator interaction surfaces of the intact VDR-RXR complex. Nat Struct Mol Biol 2011;18:556–563.
25. Helsen C, Dubois V, Verfaillie A, Young J, Trekels M, Vancraenenbroeck R, De Maeyer M, Claessens F. Evidence for DNA-binding domain–ligand-binding domain communications in the androgen receptor. Mol Cell Biol 2012;32:3033–3043.
26. Fiser A, Do R, Šali A. Modeling of loops in protein structures. Protein Sci 2000;9:1753–1773.
27. Ravikumar K, Huang W, Yang S. Coarse-grained simulations of protein-protein association: an energy landscape perspective. Biophys J 2012;103:837–845.
28. Yang S, Onuchic JN, Levine H. Effective stochastic dynamics on a protein folding energy landscape. J Chem Phys 2006;125:054910.

29. Sobolev V, Sorokine A, Prilusky J, Abola E, Edelman M. Automated analysis of interatomic contacts in proteins. Bioinformatics 1999;15: 327–332.
30. Kim YC. Hummer G. Coarse-grained models for simulations of multiprotein complexes: application to ubiquitin binding. J Mol Biol 2008;375:1416–1433.
31. Miyazawa S. Jernigan RL. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. J Mol Biol 1996;256:623–644.
32. Brooks BR, Brooks C, III, Mackerell A, Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M. CHARMM: the biomolecular simulation program. J Comput Chem 2009;30:1545–1614.
33. Yang S, Roux B. Src kinase conformational activation: thermodynamics, pathways, and mechanisms. PLoS Comput Biol 2008;4: e1000047.
34. Schlitter J, Engels M, Krger P. Targeted molecular dynamics: a new approach for searching pathways of conformational transitions. J Mol Graph 1994;12:84–89.
35. Tsui V, Case DA. Theory and applications of the generalized born solvation model in macromolecular simulations. Biopolymers 2000; 56:275–291.
36. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with namd. J Comput Chem 2005;26:1781–1802.
37. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins 2006;65:712–725.
38. Pérez A, Marchán I, Svozil D, Sponer J, Cheatham TE, III, Laughton CA, Orozco M. Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. Biophys J 2007;92:3817–3829.
39. Aaqvist J. Ion-water interaction potentials derived from free energy perturbation simulations. J Phys Chem 1990;94:8021–8024.
40. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25:1157–1174.
41. Wang J, Wang W, Kollman PA, Case DA. Automatic atom type and bond type perception in molecular mechanical calculations. J Mol Graph Model 2006;25:247–260.
42. Jakalian A, Bush BL, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. method. J Comput Chem 2000;21:132–146.
43. Jakalian A, Jack DB, Bayly CI. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. parameterization and validation. J Comput Chem 2002;23:1623–1641.
44. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, Onufriev A, Simmerling C, Wang B, Woods RJ. The Amber biomolecular simulation programs. J Comput Chem 2005;26:1688–1668.
45. Paige LA, Christensen DJ, Grn H, Norris JD, Gottlin EB, Padilla KM, Chang, CY, Ballas LM, Hamilton PT, McDonnell DP, Fowlkes DM. Estrogen receptor (ER) modulators each induce distinct conformational changes in er alpha and er beta. Proc Natl Acad Sci USA 1999;96:3999–4004.
46. Kong EH, Heldring N, Gustafsson JA, Treuter E, Hubbard RE, Pike ACW. Delineation of a unique protein-protein interaction site on the surface of the estrogen receptor. Proc Natl Acad Sci USA 2005; 102:3593–3598.
47. Wang Y, Chirgadze N, Briggs S, Khan S, Jensen E, Burris T. A second binding site for hydroxytamoxifen within the coactivator-binding groove of estrogen receptor β. Proc Natl Acad Sci USA 2006;103:9908–9911.
48. Chandra V, Huang P, Potluri N, Wu D, Kim Y, Rastinejad F. Multidomain integration in the structure of the HNF-4α nuclear receptor complex. Nature 2013;495:394–398.
49. Kuntz M, Shapiro D. Dimerizing the estrogen receptor DNA binding domain enhances binding to estrogen response elements. J Biol Chem 1997;272:27949–27956.
50. Levitt M. Nature of the protein universe. Proc Natl Acad Sci USA 2009;106:11079–11084.
51. Zheng W, Schafer NP, Davtyan A, Papoian GA, Wolynes PG. Predictive energy landscapes for protein-protein association. Proc Natl Acad Sci USA 2012;109:19244–19249.
52. Tompa P. On the supertertiary structure of proteins. Nat Chem Biol 2012;8:597–600.
53. Rochel N, Ciesielski F, Godet J, Moman E, Roessle M, Peluso-Iltis C, Moulin M, Haertlein M, Callow P, Mely Y, Svergun DI, Moras D. Common architecture of nuclear receptor heterodimers on DNA direct repeat elements with different spacings. Nat Struct Mol Biol 2011;18:564–570.
54. Orlov I, Rochel N, Moras D, Klaholz BP. Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. EMBO J 2012;31:291–300.