Recognition of LXXLL by Ligand Binding Domain of the Farnesoid X Receptor in Molecular Dynamics Simulation

Tao Zhang, Xi-Cheng Dong,* and Min-Bo Chen

Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai, P. R. China, 200032

Received March 29, 2006

The Farnesoid X receptor (FXR) has recently become a potential therapeutical target. The recruitment of coactivator protein (specified by LXXLL sequence) is the initial step in transcriptional activation of nuclear receptors (NRs). In this paper, the process of recognition of the LXXLL motif by the ligand binding domain (LBD) of FXR is observed in a 25 ns molecular dynamics simulation. The hydrophobic and hydrogen bonding interactions between the LBD and the coactivator are fully analyzed. This observation provides justification for the 'on deck' model proposed by Nettles and Greene. At last, insight to the protein−polypeptide interactions and protein conformational changes are discussed.

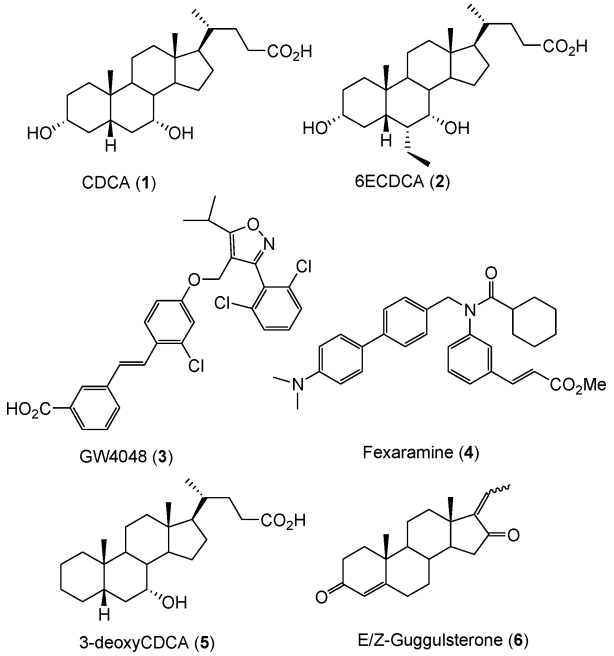
INTRODUCTION

By binding with the lipophilic steroid and retinoid hormones, the intracellular nuclear receptors (NRs) participate in a diverse range of biological activities such as development, reproduction, differentiation, and cellular homeostasis.[1,2] There are 48 transcriptional factors in the human NR family, which is divided into seven classes.[3] The Farnesoid X receptor (FXR, also called BAR, NR1H4[34]) was first observed to be activated by the farnesol at high concentration, and then the bile acids were identified as the endogenous ligand of FXR in 1999.[4−7] In the metabolic path of cholesterol, through the increase in the level of the SHP to repress the LRH-1,[8,9] FXR regulates the bile acids synthesis by a feedback repression of CYP7A1 and CYP8B1 genes encoding for cholesterol 7α-hydroxylase and 12α-hydroxylase,[10] two key enzymes in the metabolism process from cholesterol to bile acids. Now, FXR becomes an attracting potential target for treating cholesterol and bile acid related diseases.

FXR shares the classical modular architecture with other NRs,[11−14] mainly including DNA-binding domain (DBD), ligand-binding domain (LBD), and a connecting hinge region. The LBD acts as a molecular switch that controls the release of a corepressor protein and subsequent recruitment of a coactivator protein containing a highly conserved 'signature sequence' LXXLL.[15] The recruitment of the coactivator is achieved along with conformational changes of the LBD, which are usually regarded to be induced by a proper agonist. This function of NRs to associate with the coactivator is called the activation function 2 (AF2),[16] and the site for LXXLL motif is that for AF2.

The chenodeoxycholic acid (CDCA, Chart 1) **1** was identified as the endogenous ligand for FXR in 1999.[5,6,7] This discovery greatly accelerates the research of FXR. In recent years, 6-ethyl-chenodeoxycholic acid (6ECDCA,[17] Chart 1) **2**, and some nonsteroid molecules, such as GW4064[18] (Chart 1) **3** and fexaramine[19] (Chart 1) **4**, have been developed as FXR agonists. In addition, a natural product, guggulsterone[20] (Chart 1) **5**, is discovered as a special antagonist of FXR.

Chart 1. Structures of Various Modulators of FXR



Two kinds of crystal structures of LDB (PDB code: 1OSV, 1OT7,[21] and 1OSH[22]) in the presence of 6ECDCA, 3-deoxy-CDCA (Chart 1) **6**, and fexaramine, respectively, have unveiled several features of the receptor: (i) Similar to other NRs, the helix12 is packed with helix3 and helix11 forming AF2 site. (ii) The binding model of LBD with 6ECDCA differs from most endogenous ligands with their cognate receptors. (iii) Surprisingly, two coactivator fragments are found in FXR-LDB crystal structure.[21] The normal one fragment locates at the AF2 site, and the other is in a nearby position. Except for FXR and LRH-1,[23] no more than one signature sequence polypeptide is observed in the known crystal structures of NRs LBD (Figure 1).

Some studies have been made in understanding the mechanism of activating/inhibiting FXR by means of mo-

* Corresponding phone: +86 21 5492 5299; fax: +86 21 6416 6128; e-mail: dongxc@mail.sioc.ac.cn.

10.1021/ci060112v CCC: $33.50
Published on Web 10/24/2006

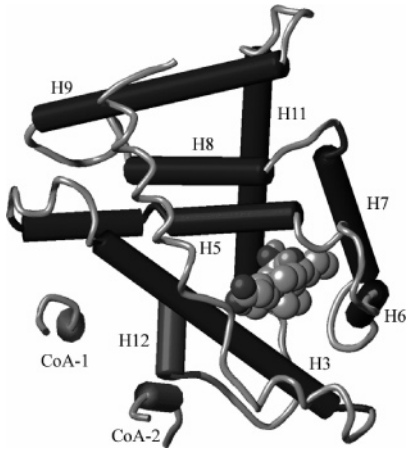


Figure 1. The crystal structure of LBD of FXR, chain B of the 1OT7. The body of LBD is represented by the black cylinder; the helix12 (P460−I465), coactivator-1 (on the AF2 site), and coactivator-2 (on the 'on deck' site) is represented by the gray cylinder; and the other part is represented by light gray tubes. The ligand is represented as a ball model. The body of LBD included the following: helix3 (A277-K300), helix7 (D356-E371), helix11 (H429-L448), helix5 (H310-F330), helix6 (D343-I349), helix8 (Q376-I387), and helix9 (R398-Y419). Note: The residues of LBD are denoted by a single-letter abbreviation and a sequence number; the residues on coactivator-1 are denoted by a three-letter abbreviation, an underline, and a sequence number; and the residues on coactivator-2 are denoted by a three-letter abbreviation and a sequence number with no sign of an underline. For example: E464, Leu_4, and Leu4, respectively.

Table 1. Details in Simulations

simulation code	elements in system	simulation time (ns)
simulation-I	LBD; 6ECDCA; coactivator-2; water (13362)	25
simulation-II	LBD; 6ECDCA; coactivator-1; water (13028)	4

lecular modeling. Nettles and Greene proposed that a midway state (meta-complex) exists in the process of LDB binding with a coactivator and that the noncanonical site for a second coactivator fragment be called the 'on deck' site.[24] Costantino and co-workers observed the stability of the helix12 and the coactivator in the presence of a bile acid agonist and clarified the relationship among LBD, coactivator, and agonist.[25] Meyer and co-workers put forward a special potential antagonistic mechanism of the guggulsterone in the recent molecular docking investigation.[26]

The 'on deck' proposal requires some direct evidences. It is unclear whether and how the coactivator enters into the canonical site from 'on deck'. Costantino and co-workers reported the release of coactivator-2 in holo-LBD.[25] However the destination of coactivator-2 is not clear. Based on the work by Nettles and Greene and Costantino et al., molecular modeling is carried out to get a better understanding for the LBD-coactivator association and find out evidence for the 'on deck' proposal in this paper.

In our work, two triunitarian models (simulation-I and simulation-II, Table 1) are designed to study the behavior of LBD-coactivator. Simulation-I contains LBD, 6ECDCA, and coactivator-2 but no coactivator-1. In this model, coactivator-2 is located at the 'on deck' site, while the helix12 is in the active position. The presence of 6ECDCA provides stability of the helix12. On the basis of the 'on deck' proposal, coactivator-2 is anticipated to move from the 'on deck' to the AF2 site. For comparison and getting some dynamics information of the LBD−coactivator−ligand complex, simulation-II is performed, in which LBD, ligand, and coactivator-1 are included.

METHODS AND MATERIALS

The initial structure of the protein was obtained from chain B of 1OT7,[21] where coactivator-1 was deleted in simulation-I, and coactivator-2 was deleted in simulation-II. Some necessary repairs for missing atoms in the residues were accomplished by the biopolymer module in Sybyl. Essential hydrogen atoms were added using pdb2gmx in Gromacs,[27] and the Gromacs force field was employed. A triclinic box was added using editconf in Gromacs, and the distance between the protein and the box was set to 12 Å. The box was filled with SPC water molecules.

The starting structure of the 6ECDCA, including its conformation and position in the ligand-binding pocket of LDB, was taken from molecular docking studies. (The docking result was almost identical to the experimental result, see the Supporting Information). The ligand topology file for Gromacs was obtained from the Dundee PRODRG2 server.[28] The charge of 6ECDCA was computed by the AM1 method using Gaussian03. Because only polar hydrogen was required by the Gromacs force field, charge and mass of nonpolar hydrogen were united to the next heavy atoms. The coordinate and topology file of the ligand were merged into those of the protein in order to complete the model building.

The strategy of simulation included minimization, restrained molecular dynamics, and unstrained molecular dynamics. The energy minimization was to remove the local strain in the peptide and to remove bad VDW contacts. The minimization was set to be converged when the maximum force was smaller than 1000 $kJ \cdot mol^{-1} \cdot nm^{-1}$. The water molecules were allowed to move freely around/inside the protein in the 20 ps restrained molecular dynamic, where all heavy atoms in the protein and the ligand were restrained with a force constant of 1000 $kcal \cdot mol^{-1} \cdot Å^{-2}$. The simulation time of unstrained molecular dynamics was 25 ns and 4 ns in simulation-I and simulation-II, respectively.

The starting temperature was set to 10 K and heated to 300 K quickly. To avoid the protein escaping from the box, the translation movement of the mass center of the protein was removed every step. The temperature and the pressure were kept constant by using Berendsen rescaling methods.[29] The protein, waters, and the ligand were put into three self-governed heat baths in order to avoid the problem of hot solvent with cold solute. The time step for integration was 2 fs. The snapshot of the system was taken every 1 ps. The particle-mesh Ewald method[30,31] was employed to calculate the long-range electrostatic interaction, and the SHAKE[32] method was used to constrain bonds involving hydrogen atoms.

For the docking procedure, the program AUTODOCK[33] was used. The Lamarckian Genetic Algorithm was adopted for docking, and the empirical binding free energy function was adopted for scoring. One hundred parallel LGA docking runs were conducted. The samples were clustered using an acceptable root-mean-square deviation (rmsd) with 2 Å tolerance.

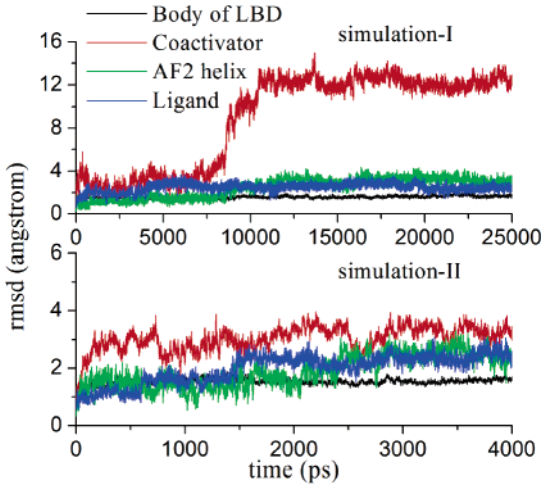


Figure 2. Variations of the rmsd of the body (black), coactivators (red), AF2 helix (green), and ligand (blue) in both simulations. The body backbone atoms of the crystal structure are taken as the reference structure in least-squares-fitting. The backbone atoms of body, coactivators, AF2 helix, and all the ligand atoms are used for measuring their rmsd values, respectively. The big increase of rmsd of coactivator-2 in simulation-I indicates the obvious position change.

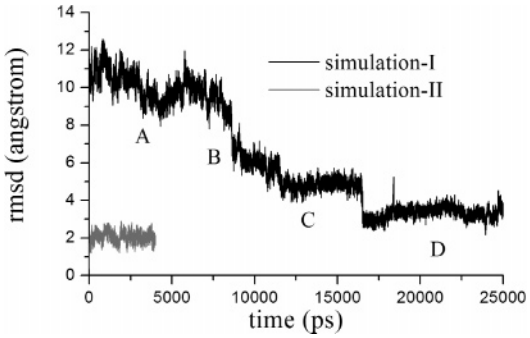


Figure 3. Variations of the rmsd of the leucine side chains of each coactivator in both simulations. The leucine side chain atoms of coactivator-1 in the crystal structure are taken as the reference. Note: One-to-one correspondence in rmsd calculation is as follows: Leu_9/Leu4, Leu_8/Leu5, Leu_5/Leu8, and Leu_4/Leu9 in simulation-I and Leu_4/Leu_4, Leu_5/Leu_5, Leu_8/Leu_8, and Leu_9/Leu_9 in simulation-II, see Figure 11a.

RESULTS

The Overall of the Simulations. FXR ligand binding domain shares the common three-layer antiparallel α helical sandwich structure with other NRs LBD. Three long helices (helix3, 7, 11) form the two outer layers of the sandwich, and the middle layer is composed of four shorter helices (helix5, 6, 8, 9). This seven helices combination is called the body of LBD in this paper. In reference to the crystal structure, the rmsd values over the backbone of the body are 1.54 ± 0.13 Å and 1.56 ± 0.15 Å in simulation-I and simulation-II, respectively (Figure 2, black lines). The low rmsd values indicate the body of LBD is very stable during the simulations.

Despite the stability of the body in simulation-I, the rmsd of coactivator-2 increases by nearly 10 Å during the middle of simulation-I from 2.96 ± 0.89 Å in the first 7.5 ns to 12.27 ± 0.34 in the last 14 ns (Figure 2). This big increase corresponds to the displacement of coactivator-2 from the 'on deck' to the AF2 site (Figure 3). With the moving of coactivator-2, the position of the AF2 helix (helix12) is also

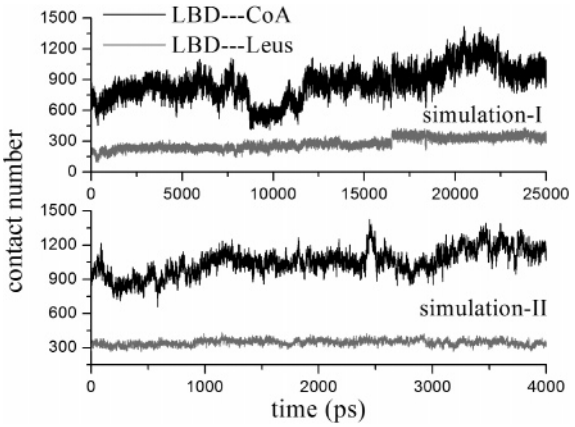


Figure 4. Variations of the contact number C in both simulations. The contact number between the coactivator and LBD, $C_{LBD-CoA}$, is represented by black curves, and the contact number between the four leucine side chains and LBD, $C_{LBD-Leus}$, is represented by gray curves. The $C_{LBD-CoA(I)}$ drops by about 20% in the middle of simulation-I, indicating partial dissociation of LBD with coactivator-2.

changed with an rmsd of 1.29 ± 0.30 Å in the first 7.5 ns and an rmsd of 3.16 ± 0.34 Å in the last 14 ns (Figures 2 and 11, see the Supporting Information for details). In simulation-II, although the higher value rmsd of coactivator-1 (3.03 ± 0.42 Å) and the AF2 helix (1.90 ± 0.57 Å) indicate a lower stability in comparison with that of the body, the hydrophobic and hydrogen bonding interaction between coactivator-1 and LBD are conserved during the whole simulation time (see the later description). The ligands are moderately stable in both simulations.

Destination of the Movement of Coactivator-2 in Simulation-I. Taking the four leucines of coactivator-1 in the crystal structure as reference, the rmsd of the four leucines are computed in both simulations (Figure 3). The leucine side chains in simulation-II are stable with an rmsd 2.03 ± 0.28 Å (Figure 3, gray). In simulation-I, the rmsd of these side chains shows the descending trend from 11 to 3 Å (Figure 3, black). This trend clearly indicates that coactivator-2 moves from the 'on deck' to the AF2 site (Figure 11). Considering the rmsd of coactivator-2 and the leucine side chains together, the movement of coactivator-2 can be divided into four stages: A (0−7.5 ns), B (7.5−12 ns), C (12−16.5 ns), and D (16.5−25 ns). The position of the side chains in coactivator-2 during stage D is stable with an rmsd 3.34 ± 0.30 Å, slightly larger than that in simulation-II (2.03 ± 0.28 Å), indicating the destination of coactivator-2 is the AF2 site.

Hydrophobic Interaction and Hydrogen Bonding. The recognition of the LXXLL motif by the NRs LBD is modulated by hydrophobic interaction and hydrogen bonding.[35] Three conserved leucines (Leu_5, Leu_8, and Leu_9 in coactivator-1) align on the face of the α-helix that packs against the hydrophobic pocket of the LBD surface. The hydrogen bonding interaction between the charged clamp (E464 and K300, Figure 9) and coactivator-1 enhances the interaction.[36]

Hydrophobic Interaction: Contact Analysis. The intensity of the hydrophobic interaction between two groups is characterized by the contact number C (Figure 4). The 'contact' is defined herein as an event in which the distance between two atoms is less than 6 Å. Based on it, the AF2

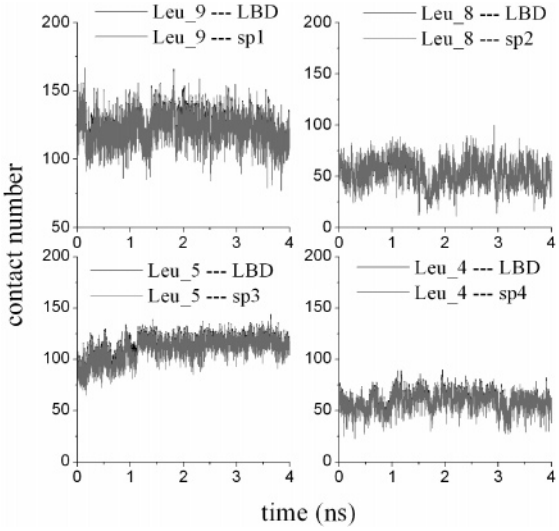


Figure 5. Variations of the contact number between a single leucine side chain and LBD (black) and that between each leucine side chain and its corresponding subpocket (gray) in simulation-II. The curves for the two cases are almost identical, indicating that each leucine always stays at its own subpocket.

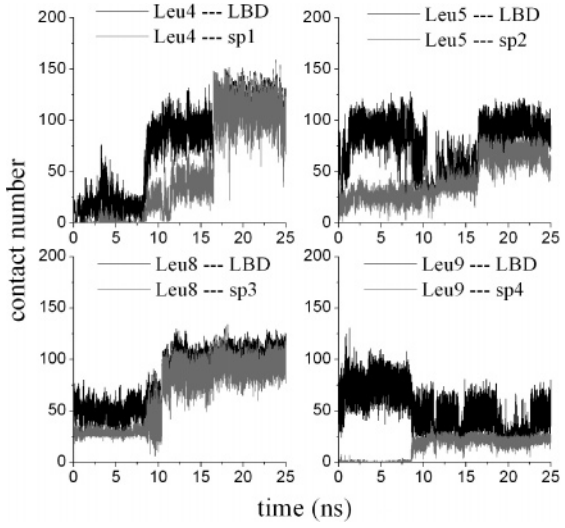


Figure 6. Variations of the contact number between LBD and single leucine side chain (black) and that between single leucine side chain and subpocket (gray) in simulation-I. Each curve forms several plateaus during simulation-I, corresponding to the four stages cited above.

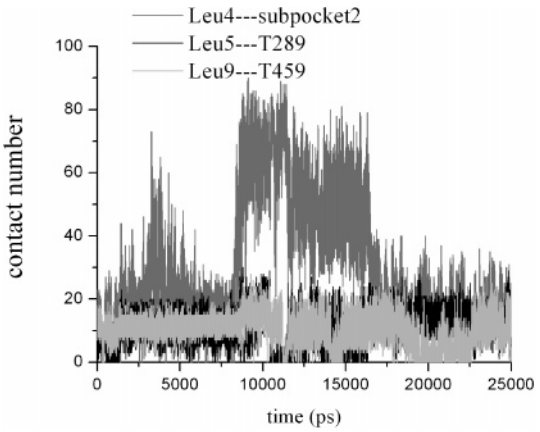


Figure 7. Variations of the contact number between Leu4 and subpocket2 ($C_{Leu4-sp2}$, gray), Leu5 and T289 ($C_{Leu5-T289}$, black), and Leu5 and T459 ($C_{Leu5-T459}$, light gray) in simulation-I. The large $C_{Leu4-sp2}$ in the middle of simulation-I indicates that Leu4 stays in subpocket2 during the period.

site in the crystal structure can be divided into four parts: subpocket1, subpocket2, subpocket3, and subpocket4 (Figure 11). The subpocket1, subpocket2, and subpocket3 recognize the three leucines in LXXLL motif. The contact number between each leucine side chain and its corresponding subpocket (C_{Leu_9-sp1}, C_{Leu_8-sp2}, C_{Leu_5-sp3}, and C_{Leu_4-sp4}) is 102, 39, 83, and 56 in chain B and 118, 34, 80, and 60 in chain A for both crystal structures.

Both $C_{LBD-CoA(II)}$ and $C_{LBD-Leus(II)}$ maintain their stability during simulation-II ($C_{LBD-CoA(II)} = 1052 \pm 116$, $C_{LBD-Leus(II)} = 347 \pm 26$). It indicates that the hydrophobic interaction between LBD and coactivator-1 is quite stable (Figure 4). For each leucine, its contact number with the corresponding subpocket is stable but is a little higher than that in the crystal structures ($C_{Leu\_9-sp1} = 122 \pm 12$, $C_{Leu\_8-sp2} = 52 \pm 13$, $C_{Leu\_5-sp3} = 111 \pm 12$, and $C_{Leu\_4-sp4} = 59 \pm 9$). These data show that the leucines always stay in their corresponding subpocket during simulation-II, and the hydrophobic interaction is conserved.

There is a 20% decrease of $C_{LBD-CoA(I)}$ in the middle of simulation-I from 801 ± 90 in stage-A (0−7.5 ns) to 667 ± 124 in stage-B (7.5−12 ns) (Figure 4). It indicates partial dissociation of the coactivator from the LBD during stage-B. Correspondingly, the contact between leucine side chains and LBD ($C_{LBD-Leus(I)}$) shows high stability during the first 16 ns ($C_{LBD-Leus(I)} = 246 \pm 33$, Figure 4), indicating a stable association of LBD with LLXXLL. In other words, the leucine side chains can be bound tightly on the LBD surface, although the whole coactivator-2 is moving quickly on the surface of the protein. The $C_{LBD-Leus(I)}$ increases dramatically by about 25% at 16 ns (Figure 4), and the rmsd of the leucine side chains decreases correspondingly (Figure 3). It declares that coactivator-2 has entered the AF2 site (see later).

Although the total contact number between LBD and four leucine side chains is stable, the contact between a single leucine side chain and LBD changes dramatically during simulation-I. In stage-A, the contact of LBD-Leus happens mainly at LBD-Leu5 and LBD-Leu9 ($C_{LBD-Leu4} = 19 \pm 10$, $C_{LBD-Leu5} = 86 \pm 17$, $C_{LBD-Leu8} = 49 \pm 8$, and $C_{LBD-Leu9} = 74 \pm 12$, Figure 6), and coactivator-2 stays at the 'on deck' site (Figure 11a).

At the beginning of stage-B, Leu4 quickly enters into the subpocket2 (Figure 7, gray), and then Leu8 steps into the center of subpocket3 (Figure 6). The motion of Leu4 initiates the whole movement of coactivator-2. At the same time, Leu5 and Leu9 are pulled out of the 'on deck' site (Figures 6 and 11b). Stage-B ends up with the preliminary allocation of Leu4 and Leu8. In stage-C, the contact for a single leucine side chain is moderately stable ($C_{LBD-Leu4} = 90 \pm 12$, $C_{LBD-Leu5} = 47 \pm 10$, $C_{LBD-Leu8} = 95 \pm 11$, and $C_{LBD-Leu9} = 42 \pm 13$). Stage-C suddenly shuts down at 16 ns, when Leu4 and Leu5 rashly move into their subpocket1 and supocket2, respectively. The total contact number of LBD-Leus acutely increases at the same time from 246 ± 33 to 341 ± 30 at 16 ns. In stage-D, Leu4 and Leu8 stay right at the predefined subpocket1 and subpocket3, respectively ($C_{LBD-Leu4} = 114 \pm 15$, $C_{LBD-Leu8} = 101 \pm 9$, $C_{Leu4-sp1} = 112 \pm 15$, and $C_{Leu8-sp3} = 95 \pm 8$). The contact between Leu5 and subpocket2 is enhanced by T289 and other residues

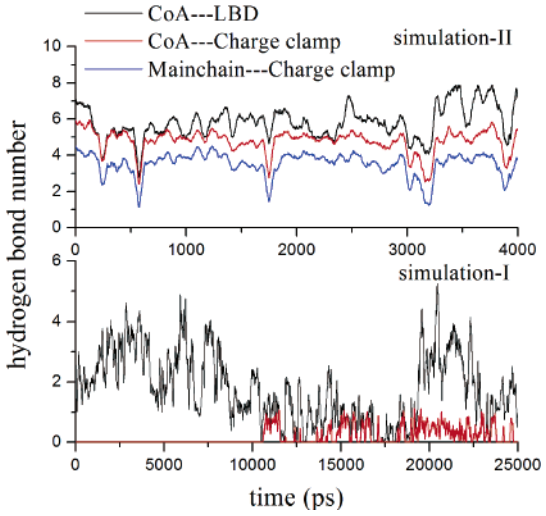


Figure 8. Variations of the number of hydrogen bonds, HB, formed between each coactivator and LBD (black, $HB_{LBD-CoA}$), the coactivator and the charge clamp (red, $HB_{charge\ clamp-CoA}$), and the coactivator's mainchain and the charge clamp (blue, $HB_{charge\ clamp-mainchain}$). The data are treated by a smoothing average of 50 adjacent points. The hydrogen bond is characterized by two features: the angle $\alpha < 30°$ formed by acceptor−donor-hydrogen and the bond length $l < 3.5$ Å between the donor and the acceptor.

($C_{LBD-Leu5} = 91 \pm 13$, $C_{Leu5-sp2} = 66 \pm 8$, and $C_{Leu5-T289} = 14 \pm 5$). It indicates Leu5 enters more deeply than the predefined subpocket2. Because of the repulsion between Asp7 and E464 (see later), no contact between Leu9 and E464 is observed in simulation-I. Thus, the $C_{Lue9-LBD}$ in simulation-I is smaller than C_{Leu_4-LBD} in simulation-II despite extra contact from T459. Actually, Leu9 is not of the member of the signal sequence LXXLL, and its behavior does not affect the recognition of LXXLL by AF2 site significantly.

From the contact analysis above, it is observed that the conserved leucines enter the AF2 site one by one. First, Leu9 enters subpocket2 at 7.5 ns, and then the Leu5 slips into subpocket3 during stage-B. Finally, Leu9 rashly switches to subpocket1, and Leu8 moves into subpocket2 simultaneously (Figures 6 and 11).

Hydrogen Bonding Interaction. Like other nuclear receptors, the 'charge clamp' consists of two highly conserved charged residues in FXR, K300 in helix3 and E464 in helix12. The hydrogen bonding interaction between the charge clamp and the coactivator's mainchain is considered as an important role in recognition of the LXXLL motif by LBD. It is suggested that recruitment of the coactivator by the receptor AF2 surface is initiated by charge interactions.[38] However, the recruitment of coactivator-2 is also observed in simulation-I although the electrostatic interaction herein is unfavorable.

The hydrogen bonding interaction between coactivator-1 and LBD is stable during simulation-II ($HB_{LBD-CoA(II)} = 5.8 \pm 1.5$), which mainly comes from the charge clamp ($HB_{charge\ clamp-CoA(II)} = 4.8 \pm 1.2$). Most of the latter is attributed to the conserved interaction between the coactivator's mainchain and the charge clamp ($HB_{charge\ clamp-mainchain(II)} = 3.5 \pm 1.1$, Figure 8). The high stability of hydrogen bonding is strengthened by the complementary electrostatic interaction. The positive K300 catches the N-terminal of LXXLL, where the negative residues (Asp10 and Asp12) are located. The negative E464 holds the C-terminal of

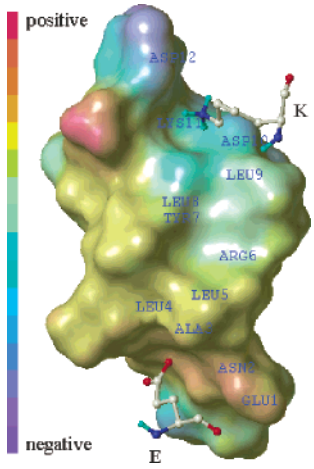


Figure 9. The Connolly surface for electrostatic potential of coactivator-1 in the crystal structure. The color ramp is from red (most positive) to purple (most negative). The charge clamp (K300 and E464) is represented by a ball-stick model. The sequence of the coactivator fragment is ENALLRYLLDKD.

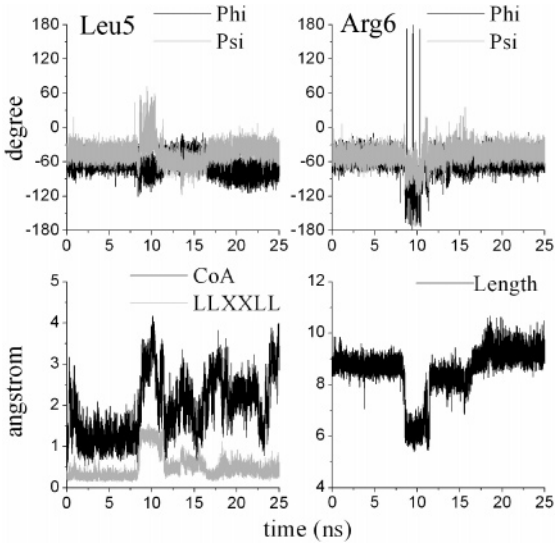


Figure 10. **a**.**b**. The variations of phi (black) and psi (gray) angles of Leu5 and Arg6. **c**. Rmsd values for backbone atoms of coactivator (black) and LLXXLL (gray) are computed. Their own starting structures are taken as the references. **d**. The distance between the alpha carbons of Leu4 and Leu9 is computed, representing the length of LLXXLL.

LXXLL, where the polar Asn2 is located. Figure 9, the Connolly surface[38] for the electrostatic potential of coactivator-1, clearly shows this complement.

In simulation-I, the hydrogen bonding interaction between coactivator-2 and LBD ($HB_{LBD-CoA(I)} = 1.8 \pm 1.4$) is much weaker than that in simulation-II. The hydrogen bonding interaction between the charge clamp and coactivator-2 is still weak ($HB_{charge\ clamp-CoA(I)} = 0.3 \pm 0.5$ in stage-D), and no hydrogen bond between the charge clamp and the mainchain of coactivator-2 is observed, even after the LXXLL enters the AF2 site completely. The hydrogen bonding described above is not important for recognition of LXXLL by the AF2 site in simulation-I, as compared to that in simulation-II.

This weak hydrogen bonding interaction comes from the 'opposite' binding direction of LXXLL. As seen in simulation-II, the electrostatic interaction is favorable. The two

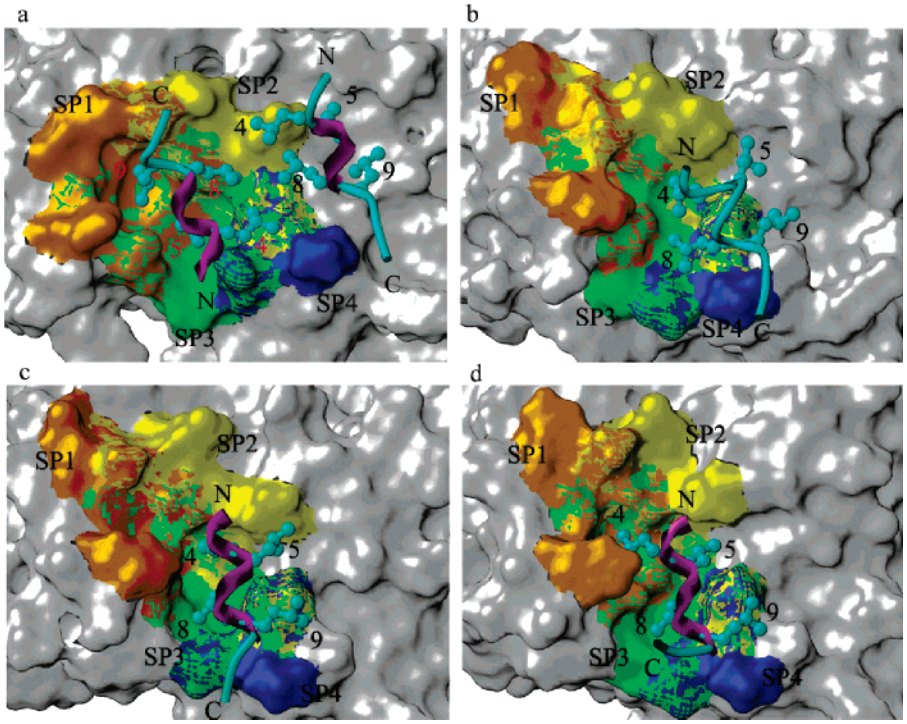


Figure 11. The Connolly surfaces of the AF2 site at different stages in simulation-I. Based on the contact between the single leucine side chain and LBD, the AF2 site is divided into four parts: subpocket1 (for Leu_9, in orange), subpocket2 (for Leu_8, in yellow), subpocket3 (for Leu_5, in green), and subpocket4 (Leu_4, in blue). The subpocket1 includes V296, T299, K300, F305, Q306, A310, Q313, I314, and L317; the subpocket2 includes V292, Q293, V296, E297, K300, and L461; the subpocket3 includes V292, V296, Q313, I314, L317, K318, L461, E464, and I465; and the subpocket4 includes P460, L461, and E464. The coactivators are represented by a magenta ribbon (for helix) and a cyan tube (for loop), and the leucine side chains are represented by the cyan ball-stick model. **a**: LBD, coactivator-1, and coactivator-2 of 1OT7 (coactivaor-1 is not included in simulations-I). **b**−**d**: snapshots at 9, 14, and 25 ns in simulation-I, respectively. SP stands for subpocket.

coactivators are antiparallelly aligned in crystal structure (Figure 11a). So, our starting structure for coactivator-2 in simulation-I is in the opposite direction compared to the case of coactivator-1 in simulation-II. This 'opposite' direction leads to an incompatible electrostatic interaction and prevents the formation of the conserved hydrogen bonds.

The Conformational Change of Coactivator-2 in Simulation-I. In stage-B when coactivator-2 walks from the 'on deck' to the AF2 site, both the hydrophobic and the hydrogen bonding interactions decrease, indicating partial dissociation of coactivator-2 from LBD. For coactivator-2 itself, the intraconformation of LLXXLL changes a great deal, rather than a two-turn helix in stage-A. The most obvious change takes place in the dihedrals of Lue5 and Arg6 backbone (Figure 10a,b), shortening the length of LLXXLL from 8.76 $\pm$ 0.23 Å in stage-A to 6.62 $\pm$ 0.91 Å in stage-B. After coactivator-2 enters the AF2 site, the whole coactivator suffers a wide fluctuation in rmsd 2.16 $\pm$ 0.58 Å (Figure 11c, black) because of incompatible electrostatic interaction.

Summary of the Movement of Coactivator-2 in Simulation-I. The whole process of recognition of coactivator-2 by the AF2 site in LBD can be better understood using Figure 11. At the beginning of simulation-I (stage-A), coactivator-2 stays at the 'on deck' site, while Leu4 and Leu8 wander at the edge of subpocket2 and subpocket3 (yellow and green in Figure 11a). The affinity between Leu4 and subpocket2 initiates the movement of coactivator-2, and then attraction between Leu8 and subpocket3 pulls the whole coactivator-2 into the AF2 site (Figures 2, 3, 6, and 11b). Therefore, the two leucine side chains (Leu4 and Leu7) are recognized by the AF2 site (Figure 11c). Finally, the Leu4 side chain switches into subpocket1, and the Leu5 enters the center of subpocket2 in 1 ps. This very quick switch accomplishes the recognition of LXXLL by LBD (Figure 11d).

DISCUSSION

About Simulation-I. Using the molecular dynamics simulation, the process of recognition of LXXLL by LBD is observed. The simulation shows that the three leucines of LXXLL are recognized by the AF2 site consecutively, although the charge clamp cannot catch coactivator-2 because of an incompatible electrostatic interaction. We consider that it might take hundreds of nanoseconds (even more) to adjust this incompatible electrostatic interaction if simulation-I is elongated.

In the real process of the recognition, it is considered that LXXLL would enter the place in the right charge direction because the electrostatic effect plays a main role in the long-range interaction.[16] In other words, the charge clamp first recognizes the direction of the coactivator, and then the AF2 site recognizes the specific leucine side chains of LXXLL. Fortunately, the coactivator that is used in our model includes an LLXXLL sequence, and the coactivator is close to the AF2 site. So, the AF2 site can still recognize the side chains of LLXXL despite its wrong direction. Of course, the movement of coactivator-2 also indicates that the hydrophobic interaction is stronger than the electrostatic interaction in the recognition of LXXLL by the AF2 site.

In addition, the construction of the initial model of simulation-I in our paper is based on the 'on deck' proposal,

and the consequent result of the simulation gives the details of the process predicted by the proposal. Hence, the simulation lays down the justification for the 'on deck' proposal.

Quick and Slow Process in Molecular Recognition. The transition time of Leu4 from the outside to the inside of subpocket2 is about 0.5 ns, but the wandering time for Leu4 waiting for a chance of recognition by subpocket2 is about 8 ns (Figures 6, 7, and 11). The transition from stage-C to stage-D looks very strange at 16 ns (Figures 3 and 6). Only based on the data from the first 16 ns, one cannot predict when or even whether the transition will happen. The transition time of Lue4 entering subpocket1 is short, about 1 ps, but its wandering time for waiting the chance of recognition by subpocket1 is about 4000 ps. A similar phenomenon is also observed in protein folding studies. The wandering time for the miniprotein beta3s to wait for the folding event is much longer than that to do folding itself.[37]

The transition time is much shorter than the corresponding wandering time. During the transition time, the process is dominated by a strong interaction rather than the thermal motion of molecules. It happens only in the moment when the geometric match is satisfied. However, the process during the wandering time depends only on weak interactions, e.g. the thermal motion of nuclear skeleton of molecules. The wandering process is thus stochastic in nature. In comparison, the process of transition is more deterministic and predictable (equivalent to less stochastic) than the wandering process. For example, after 4 ns waiting for the chance of conformational match, the event of transition for Leu4 entering subpocket1 is realized only within 1 ps.

Why Is the Motion of Coactivator-2 so Quick? The characteristic time of the rigid-body (helix) motion of protein is usually considered to be $10^{-9}-10^{-6}$ s.[39] However, coactivator-2 moves from the 'on deck' to the AF2 site only within 20 ns in simulation-I. Why is the motion so quick? Coactivator-1 in the crystal structure is directly deleted for constructing the initial model of simulation-I. It means the AF2 site is of the shape for accepting three leucine side chains at the beginning of the simulation. The Leu4 is already very close to subpocket2, and it is a benefit for Leu4 to enter into subpocket2. Both of these two effects reduce the wandering time for the conformational matches between Leu4 and subpocket2 and between Leu8 and subpocket3. These two conformational matches drive the entire coactivator-2 to the destination. From the viewpoint of potential energy surface (PES), the starting conformation of the complex (LBD−coactivator2−ligand) is on the slope not at the minimum of the PES, that makes the complex easier to overcome the energy barrier.

Time Scale for Protein Motion. Our simulation reveals the time scale of three kinds of protein motion.

The transition of Leu4 from subpocket2 to subpocket1 only lasts 1 ps, indicating that a strong recognition of Leu4 by subpocket1 can be finished within the time scale of a picosecond.

Leu4 fails entering subpocket2 at the moment of 3.5 ns for an inappropriate position of Leu7 (Figures 6 and 7), but the event of the coactivator's movement due to a synergetic conformational match of Leu4 and Leu7 happens after an additional 4 ns of wandering. Herein, we might estimate that the wandering time for the synergetic match is about 8 ns. But these data are only for this simulation because the process is dominated by stochastic thermal motion.

There exists a third kind of motion between them, called site-oriented wandering. For example, it takes about 4 ns for coactivator-2 to walk from the 'on deck' to the AF2 site, about 10 Å in distance. It indicates that the site-oriented wandering of polypeptide from one site of protein to another can be achieved within the time scale of a nanosecond.

For an individual molecular dynamics simulation, one does not know when a conformational match occurs; even if you know, there do exist some favorable conformations. It means that the wandering time for conformational change cannot be estimated based on data from a few simulations in contrast to prediction of the transition time.

CONCLUSION

The recognition of LXXLL by LBD was observed in simulation-I despite unfavorable electrostatic interaction. The detailed analysis of the hydrophobic interaction by counting the contact number and the hydrogen bonding interaction shows a clear picture for movement of the coactivator from the 'on deck' to the AF2 site. This observation provides justification for the 'on deck' proposal. In addition, based on some interesting phenomena observed in the simulations, the protein−polypeptide interaction and protein conformational changes are discussed.

Supporting Information Available: Docking result of 6ECDCA, structural comparison between two crystal structures, and the conformational change of helix12 in simulation-I. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Gronemeyer, H.; Gustafsson, J-Å.; Laudet, V. Principles for Modulation of the Nuclear Receptor Superfamily. *Nat. Rev. Drug Discov.* **2004**, *3*, 950−964.
(2) Francis, G. A.; Fayard, E.; Picard, F.; Auwerx, J. Nuclear Receptors and the Control of Metabolism. *Annu. Rev. Physiol.* **2003**, *65*, 261−311.
(3) Rechavi, M. R.; Garcia, H. E.; Laudet, V. The Nuclear Receptor Superfamily. *J. Cell. Sci.* **2003**, *116*, 585−586.
(4) Pellicciari, R.; Costantino G.; Fiorucci S. Farnesoid X Receptor: From Structure to Potential Clinical Applications. *J. Med. Chem.* **2005**, *48*, 5383−5403.
(5) Wang, H.; Chen, J.; Hollister, K.; Sowers, L. C.; Forman, B. M. Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. *Mol. Cell* **1999**, *3*, 543−553.
(6) Park, D. J.; Blanchard, S. G.; Bledsoe, R. K.; Chandra, G.; Consler, T. G.; Kliewer, S. A.; Stimmel, J. B.; Willson, T. M.; Zavacki, A. M.; Moore, D. D.; Lehmann, J. M. Bile acids: natural ligands for and orphan nuclear receptor. *Science* **1999**, *284*, 1356−1368.
(7) Makishima, M.; Okamoto, A. Y.; Repa, J. J.; Tu, H.; Learned, R. M.; Luk, A.; Hull, M. V.; Lustig, K. D.; Mangelsdorf, D. J.; Shan, B. Identification of a nuclear receptor for bile acids. *Science* **1999**, *284*, 1362−1365.
(8) Goodwin, B.; Jones, S. A.; Price, R. R.; Watson, M. A; McKee, D. D.; Moore, L. B.; Galardi, C.; Wilson, J. G.; Lewis, M. C.; Roth, M. E.; Maloney, P. R.; Willson, T. M.; Kliewer, S. A. A regulatory cascade of the nuclear receptors FXR, SHP-1, and LRH-1 repersses bile acid biosynthesis. *Mol. Cell* **2000**, *6*, 517−526.
(9) Lu, T. T.; Makishima, M.; Repa, J. J.; Schoonjans, K.; Kerr, T. A.; Auwerx, J.; Mangelsdorf, D. J. Molecular basis for feedback regulation of bile acid synthesis by nuclear receptors. *Mol. Cell* **2000**, *6*, 507−515.
(10) Chen. W.; Owsley, E.; Yang, Y.; Stroup, D.; Chiang, J. Y. Nuclear receptor-mediated repression of human cholesterol 7alpha-hydroxylase gene transcription by bile acids. *J. Lipid Res.* **2001**, *42*, 402−412.
(11) Carson-Jurica, M. A.; Schrader, W. T.; O'Malley, B. W. Steroid receptor family: structure and fuction. *Endocr. Rev.* **1990**, *11*, 201−202.

(12) Wurtz, J. M.; Bourguet, W.; Renaud, J. P.; Vivat, V.; Chambon, P.; Moras, D.; Gronemeyer, H. A canonical structure for the ligand-binding domain of nuclear receptors. *Nat. Struct. Biol.* **1996**, *3*, 87−94.
(13) Ingraham, H. A.; Redinbo, M. R. Orphan nuclear receptors adopted by crystallography. *Curr. Opin. Struct. Biol.* **2005**, *15*, 708−715.
(14) Li, Y.; Lambert, M. H.; Xu, H. E. Activation of Nuclear Receptors: A Perspective from Structural Genomics. *Structure* **2003**, *11*, 741−746.
(15) Heery, D. M.; Kalkhoven, E.; Hoars, S.; Parker, M. G. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **1997**, *387*, 733−736.
(16) Warnmark, A.; Treuter, E.; Wright, A. P. H.; Gustafsson, J.-A. Activation Function 1 and 2 of Nuclear Receptors: Molecular Strategies for Transcriptional Activation. *Mol. Endocrinol.* **2003**, *17*, 1901−1909.
(17) Pellicciari, R,; Fiorucci, S.; Camaioni, E.; Clerici, C.; Costantino, G.; Maloney, P. R.; Morelli, A.; Parks, D, J.; Willson, T. M. 6alpha-ethyl-chenodeoxycholic acid (6-ECDCA), a potent and selective FXR agonist endowed wiht anticholestatic activity. *J. Med. Chem.* **2002**, *45*, 3569−3572.
(18) Maloney, P. R.; Parks, D. J.; Haffner, C. D.; Fivush, A. M.; Chandra, G.; Plunket, K. D.; Creech, K. L.; Moore, L. B.; Wilson, J. G.; Lewis, M. C. Identification of a chemical tool for the orphan nuclear receptor FXR. *J. Med. Chem.* **2000**, *43*, 2971−2974.
(19) Nicolaou, K, C.; Evans, R. M.; Roecker, A, J.; Hughes, R.; Downes, M.; Pfefferkorn, J. A. Discovery and optimization of nonsteroidal FXR agonists from natural product-like libraries. *Org. Biomol. Chem.* **2003**, *1*, 1079−1092.
(20) Urizar, N. L.; Liverman, A. B.; Dodds, D. T.; Silva, F. V.; Ordentlich, P.; Yan, Y. et al. A natural product that lowers cholesterol as an antagonist ligand for FXR. *Science* **2002**, *296*, 1703−1706.
(21) Mi, L. Z.; Devarakonda, S.; Harp, J. M.; Hand, Q.; Pellicciari, R.; Willson, T. M.; Khorasanizadeh, S.; Rastinejad, F. Structural basis for bile acid binding and activation of the nuclear receptor FXR. *Mol. Cell* **2003**, *11*, 1093−1100.
(22) Downes, M.; Verdecia, M. A.; Roecker, A. J.; Hughes, R.; Hogenesc, J. B.; Kast-Woelbern, H. R.; Bowman, M. E.; Ferrer, J. L.; Anisfeld, A. M.; Edwards, P. A.; Rosenfeld, J. M.; Alvarez, J. G.; Noel, J. P.; Nicolaou, K. C.; Evans, R. M. A chemical, genetic, and structural analysis of the nuclear bile acid receptor FXR. *Mol. Cell* **2003**, *11*, 1079−1092.
(23) Wang, W.; Zhang, C.; Marimuthu, A.; Krupka, H. I.; Tabrizizad, M.; Shellore, R.; Mehra, U.; Eng, K.; Nguyen, H.; Settachatgul, C. The crystal structures of human steroidogenic factor-1 and liver receptor homologue-1. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7505−7510.
(24) Nettles, K. W.; Greene, G. L. Nuclear receptor ligands and cofactor recruitment: is there a coactivator "'on deck'"? *Mol. Cell* **2003**, *11*, 850−851.
(25) Costantino, G.; Entrena-Guadix, A.; Macchiarulo, A.; Gioiello, A.; Pellicciari, R. MDS of the ligand binding domain of farnesoid x receptor (FXR). Insight into helix-12 stability and coactivator peptide stabilization in response to agonist binding. *J. Med. Chem.* **2005**, *48*, 3251−3259.
(26) Meyer, U.; Costantino, G.; Macchiarulo, A.; Pellicciari, R. Is anagonism of E/Z-guggulsterone at the farnesoid x receptor mediated by a noncanonical bind site? A molecular modeling study. *J. Med. Chem.* **2005**, *48*, 6948−6955.
(27) Van der Spoel, D.; Lindahl, E.; Hess, E.; Groenhof, G.; Mark, A. E.; Berrendsen, H. J. C. Gromacs: fast, flexible and free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(28) Schuettelkopf, A. W.; van Aalten, M. F. PRODRG − a tool for high-throughput crystallography of protein−ligand complexes. *Acta Crystallogr.* **2005**, *60*, 1355−1363.
(29) Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(30) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N-log-(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(31) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh ewald potential. *J. Chem. Phys.* **1995**, *103*. 8577−8592.
(32) Ryckaert, J. P.; Ciccotti, G.; Berendesn, H. J. C. Numerical intergration of the Cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(33) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian Genetic Algorithm and empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.
(34) Nuclear Receptors Nomenclature Committee. A Unified Nomenclature System for the Nuclear Receptor Superfamily. *Cell* **1999**, *97*, 161−163.
(35) Savkur, R. S.; Burris, T. P. The coactivator LXXLL nuclear receptor recognition motif. *J. Peptide. Res.* **2004**, *63*, 207−212.
(36) Bin, H.; Wilson, E. M. Electrostatic Modulation in Steroid Receptor Recruitment of LXXLL and FXXLF motifs. *Mol. Cell Biol.* **2003**, *23*, 2135−2150.
(37) Cavalli, A.; Haberthur, U.; Paci, E.; Caflisch, A. Fast protein folding on downhill energy landscape. *Protein Sci.* **2003**, *12*, 1801−1803.
(38) Connolly, M. L. Analytical molecular surface calculation. *J. Appl. Crystallogr.* **1983**, *16*, 548−558
(39) Adcock, S. A.; McCammon, J. A. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.* **2006**, *106*, 1589−1615.

CI060112V