PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

The *Caenorhabditis elegans* DAF-12 nuclear receptor: Structure, dynamics, and interaction with ligands

Lautaro D. Alvarez,[1,2] Pau Arroyo Mañez,[1,3] Darío A. Estrin,[3] and Gerardo Burton[1]*

[1] Departamento de Química Orgánica and UMYMFOR (CONICET-UBA), Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina

[2] IFIBYNE (CONICET-UBA), Argentina

[3] Departamento de Química Inorgánica, Analítica y Química Física and INQUIMAE-CONICET. Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Argentina

ABSTRACT

A structure for the ligand binding domain (LBD) of the DAF-12 receptor from *Caenorhabditis elegans* was obtained from the X-ray crystal structure of the receptor LBD from *Strongyloides stercoralis* bound to (25*R*)-Δ^7-dafachronic acid (DA) (pdb:3GYU). The model was constructed in the presence of the ligand using a combination of Modeller, Autodock, and molecular dynamics (MD) programs, and then its dynamical behavior was studied by MD. A strong ligand binding mode (LBM) was found, with the three arginines in the ligand binding pocket (LBP) contacting the C-26 carboxylate group of the DA. The quality of the *ce*DAF-12 model was then evaluated by constructing several ligand systems for which the experimental activity is known. Thus, the dynamical behavior of the *ce*DAF-12 complex with the more active (25*S*)-Δ^7-DA showed two distinct binding modes, one of them being energetically more favorable compared with the 25*R* isomer. Then the effect of the Arg564Cys and Arg598Met mutations on the (25*R*)-Δ^7-DA binding was analyzed. The MD simulations showed that in the first case the complex was unstable, consistent with the lack of transactivation activity of (25*R*)-Δ^7-DA in this mutant. Instead, in the case of the Arg598Met mutant, known to produce a partial loss of activity, our model predicted smaller effects on the LBM with a more stable MD trajectory. The model also showed that removal of the C-25 methyl does not impede the simultaneous strong interaction of the carboxylate with the three arginines, predicting that 27-nor-DAs are putative *ce*DAF-12 ligands.

Proteins 2012; 80:1798–1809.

Key words: dafachronic acid; *C. elegans;* molecular dynamics; nuclear receptor, DAF-12.

INTRODUCTION

In the last decades, the nematode worm *Caenorhabditis elegans* (*ce*) has aroused interest as an important model organism for biomedical research, particularly in the functional characterization of novel drug targets.[1] In this context, the study of the nuclear receptor (NR) *ce*DAF-12, partially due to the discovery of its ligands the dafachronic acids (DAs),[2] has acquired special attention because it is involved in multiple physiological functions, such as developmental timing, metabolism, fertility, and longevity.[3] DAF-12 is a modular protein which acts as a transcription factor induced by ligands, modulating the expression of genes involved in the nematode life cycle.[4] Numerous studies have shown that DAF-12 constitutes a main switch in the signaling pathway controlling dauer formation, a nonfeeding, nonreproducing, and stress-resistant state, which is specially adapted for long-term survival.[5,6] Thus, in worms under unfavorable environmental conditions due to excessive heat, crowding, or lack of food, DAs are not biosynthesized, and subsequently the DAF-12/corepressors complexes initiate the dauer larvae formation program. In contrast, in the presence of DAs the complexes

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

Abbreviations: DA, dafachronic acid; GAFF, general AMBER force field; LBD, ligand binding domain; LBM, ligand binding mode; LBP, ligand binding pocket; MD, molecular dynamics; MSMD, multiple steered molecular dynamics; NR, nuclear receptor.

Grant sponsor: Agencia Nacional de Promoción Científica y Tecnológica; Grant number: PICT2010-0623; Grant sponsor: Universidad de Buenos Aires; Grant number: 20020100100281.

*Correspondence to: Gerardo Burton; Departamento de Química Orgánica, Facultad de Ciencias Exactas y Naturales, Pabellón 2, Ciudad Universitaria, C1428EGA Ciudad de Buenos Aires, Argentina. E-mail: burton@qo.fcen.uba.ar.

Received 2 January 2012; Revised 13 March 2012; Accepted 18 March 2012
Published online 6 April 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24076

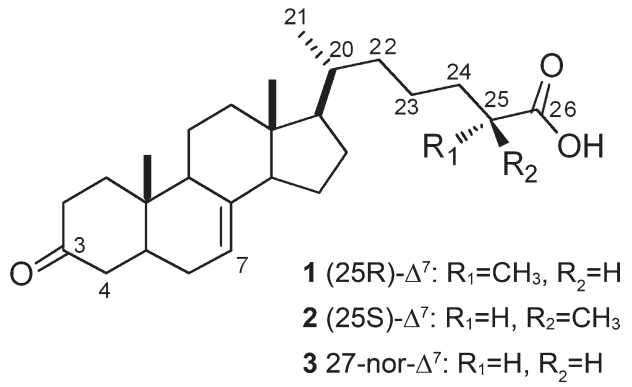

Figure 1

Structure of the Δ^7-DAs (**1**, **2**) and the 27-nor analog (**3**)

recruit coactivators and activate the transcription of genes that drive the nematode to normal reproductive development. In addition to this developmental role, DAF-12 also plays important functions in adult longevity. Parasitic nematodes such as *Strongyloides stercoralis* (ss) or *Ancylostoma caninum* also have a DAF-12 ortholog implicated in essential development processes.[7]

DAs are C-27 cholesterol metabolites biosynthesized by several enzymes of the *daf* family, including the cytochrome P450 oxidase DAF-9 which introduces a carboxylic acid group in the C-26 position (Fig. 1). Moreover, the DAs have a 3-keto group and an unsaturated double bond at C-7 or C-4 of the steroid nucleus. Δ^7 DAs have been shown to be more active and abundant in the animal. Remarkably, the configuration of the C-25 methyl group is an important determinant on complex ligand–receptor activity. Thus, transactivation assays and *in vivo* experiments have shown that the 25*S* DAs are more active than their corresponding 25*R* diastereoisomers.[8–10] Although the resulting effect of the NR depends on several factors such as receptor number and complex ability for recruiting coactivators, usually the ligand affinity is the main determinant on the transactivation activity–dose curves. Therefore, a precise characterization of the molecular basis of DAs binding mode could considerably improve our understanding on DAF-12 action.

The crystal structure of the ligand-binding domain (LBD) of *ce*DAF-12 has not been reported yet, but the *ss*DAF-12 crystal structure has been solved recently bound to the (25*R*)-Δ^7-DA (pdb:3GYU) and (25*R*)-Δ^4 ligands (pdb:3GYT).[7] Although the sequence similarity between *ss*DAF-12 and *ce*DAF-12 LBDs is 42%, this percentage increases to more than 70% when only the ligand binding pocket (LBP) residues are considered, suggesting a similar binding mode for the DAs. The X-ray crystal structure of the *ss*DAF-12 shows that the receptor is formed by 12 α-helices and three β-strands folded into the canonical NR architecture, with the ligand immersed in a cavity surrounded by residues of helices H2, H3, H5, and H7 [Fig. 2(a)]. The LBP of *ss*DAF-12 consists of 28 residues, mainly nonpolar, that make contact with the hydrophobic steroid carbon skeleton and a few polar residues which maintain electrostatic interactions with the ligand oxygen atoms [Fig. 2(b)]. In both (25*R*)-Δ^4 and (25*R*)-Δ^7-DA complexes, the steroid is oriented with the 3-keto group toward the Gln637 and the C-26 carboxylate group close to Thr562, Thr613, and Arg599 residues. The strong interaction between the negative charge density of the ligand's carboxylate group and the positive charge density of Arg599 guanidinium group appears to be especially important to ligand–receptor recognition. Besides the ligand, Arg599 interacts with Asn533, a H1 residue located in the surface of the protein which separates the LBP from the solution. The relevance of Arg599 in ligand binding is confirmed by the fact that the Arg599Met and Arg599Lys mutations eliminate completely the *ss*DAF-12/(25*R*)-Δ^7-DA activity.[7] Another arginine (Arg565) forming part of the LBP is also close to the C-26 carboxylate group; however, the crystal structures revealed that this residue does not contact the ligand but is oriented toward the H1–H2 loop, interacting with their oxygen backbone atoms. Nevertheless, since mutational studies have shown that the Arg565Cys mutant of *ss*DAF-12 is also unable to transactivate reporter genes,[7] its precise role on DAs action should be analyzed, considering the possibility that this residue is also implicated in ligand recognition.

A comparative analysis of the sequences corresponding to both LBPs shows that in the *ce*DAF-12 receptor the C-26 carboxylate group should be surrounded by three arginine residues instead of two [Fig. 2(c)]. Thus, besides the conserved Arg598 (Arg599 in *ss*DAF-12 numbering) and Arg564 (Arg565), the substitution of Val603 to Arg602 adds an additional point of positive charge density to the recognition of the acid ligand group. This substitution is accompanied by the substitution of the neutral Asn533 in *ss*DAF-12 by the negatively charged Asp532, thus conserving the overall net charge of the system. Like in the *ss*DAF-12 receptor, the Arg564Cys mutation abolished completely the *ce*DAF-12 activity[2,7]; instead, the Arg598Met mutation conserves ~60% of the wild type receptor activity,[7] suggesting that the additional Arg602 might actually play an active role in ligand binding.

Here, with the purpose of obtaining initial coordinates to be used in molecular modeling studies, we constructed a model of the *ce*DAF-12 receptor from the *ss*DAF-12 crystal coordinates. In particular, we were interested in the study of the ligand-binding mode (LBM), since a detailed knowledge would facilitate the rational design of novel *ce*DAF-12 ligands. The model was constructed in the presence of the (25*R*)-Δ^7-DA using a combination of Modeller, Autodock, and molecular dynamics (MD) programs, and then the dynamical behavior of the complex was studied by MD. To evaluate the quality of the *ce*DAF-12 model, we studied the binding mode of the more active (25*S*)-Δ^7-DA, and the effect of the Arg564Cys and Arg598Met muta-

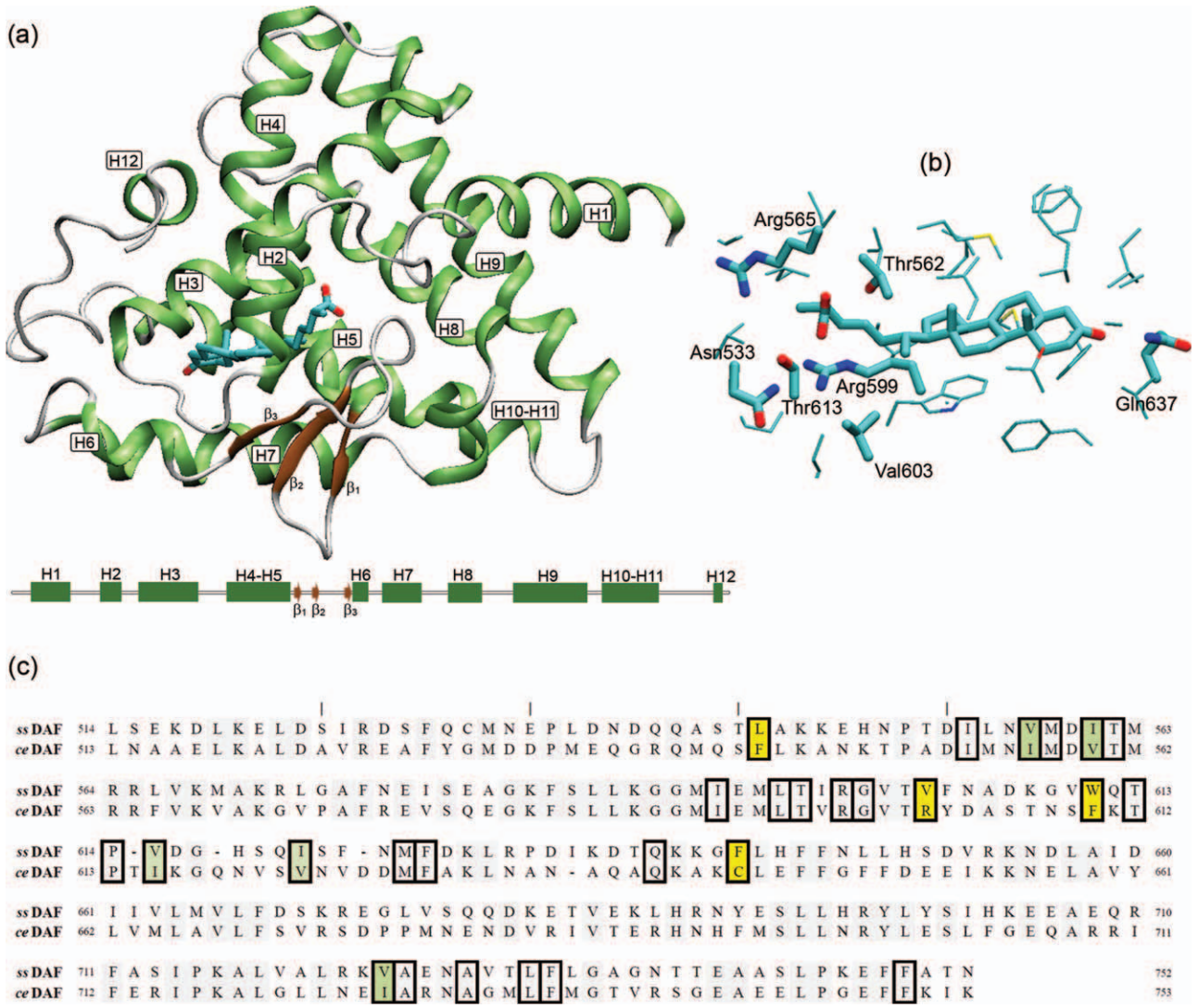

Figure 2

(**a**) Global structure and secondary sequence of the *ss*DAF-12/(25*R*)-Δ^7-DA complex (pdb:3GYU). (**b**) Detailed view of the LBP. (**c**) Sequence alignment of *ss*DAF-12 and *ce*DAF-12 receptors performed with the Modeller program. The grey shadows indicate conserved residues and the boxes indicate the residues that form part of the LBP. The green shadows and the yellow shadows indicate minor and major changes of LBP's residues, respectively.

tions on the (25*R*)-Δ^7-DA binding. Finally, for comparative purposes, the dynamical behavior of the *ss*DAF-12/(25*R*-Δ^7-DA) complex was analyzed. Taken together, the results obtained indicate that our *ce*DAF-12 model may be used to explain a number of experimental results, thus constituting a valuable template to investigate the *ce*DAF-12 molecular mechanism of action.

METHODS

Initial structures of ceDAF-12 and ssDAF-12 complexes

The initial structure of the *ce*DAF-12/(25*R*)-Δ^7-DA complex was obtained using a combination of Modeller 9 v8[11] and Autodock 4.2[12] programs. With Modeller, the LBD sequences of *ss*DAF-12 and *ce*DAF-12 receptors were aligned [Fig. 2(c)], and then three models of the *ce*DAF-12 receptor were obtained from the crystal coordinates of the *ss*DAF-12/(25*R*)-Δ^7-DA complex (pdb:3GYU). The best ranked model produced by Modeller was used to dock (25*R*)-Δ^7-DA into the *ce*DAF-12 receptor. The Autodock 4.2 method was applied considering as rotatables the torsion angles of the steroid side chain and the Arg602 side chain. To optimize the results, the C16-C17-C20-C21 and the C17-C20-C22-C23 torsion angles of the steroid were kept fixed to the values observed in the crystal structure (pdb:3GYU). A grid of 46 × 46 × 60 points with a spacing of 0.2Å centered in the LBP was calculated and used to obtain 200 runs of the genetic algorithm method. Finally, the best solution of the

most probable cluster in which the Arg602 points to the LBP was selected to perform the MD simulation of the *ce*DAF-12/(25*R*)-Δ^7-DA complex (System 1).

The initial structure of the other *ce*DAF-12 systems (Systems 2–5) were obtained from the 10 ns receptor coordinates of the System 1 trajectory. In the case of (25*S*)-Δ^7-DA (System 2), the complex was constructed with Autodock 4.2 with the same parameters used above, except that the Arg602 side chain was considered fixed. In the case of the 27-nor-Δ^7-DA steroid (**3**) (System 3), the complex was obtained by simple deletion of the C-25 methyl group (that results in an automatic replacement by hydrogen by the Tleap module of Amber 11). The *ce*DAF-12 mutant complexes (Systems 4 and 5) were constructed deleting the Arg564 or Arg598 side chain and introducing the cysteine or methionine side chains, respectively, with the Tleap module of Amber 11.[13]

In the case of the *ss*DAF-12 receptor (System 6), the crystal structure (pdb:3GYU) was used as initial structure for the MD simulation. However, as a non stable trajectory was obtained, and taking into account that the assignment of asparagine and glutamine residues in crystal structures is not trivial, we analyzed the Asn533 assignment with the NQ-Flipper program,[14] a web service based on mean force potentials to automatically detect and correct erroneous rotamers. We found that the NQ-Flipper preferred rotamer differed from the crystal assignment, and in view of this we used the NQ-Flipper corrected structure to carry out the MD simulation.

In all cases, to build the corresponding force field parameters of the ligand RESP (restraint electrostatic potential), atomic partial charges were computed using the HF method with the 6-31G** basis set in the quantum chemistry program Gaussian 03.33[15] for the corresponding HF-optimized structures, following the standard procedure of the Amber force field.

MD simulations

MD of systems 1–6 were performed with the AMBER 11 software package.[13] The ligand parameters were assigned according to the general AMBER force field (GAFF) and the corresponding RESP (restraint electrostatic potential) atomic partial charges using the Antechamber. The Amber99 force field parameters were used for all receptor residues.[16] The complexes were immersed in an octahedral box of TIP3P water molecules using the Tleap module, giving final systems of around 26,000 atoms. The systems were initially optimized and then gradually heated to a final temperature of 300 K. Starting from these equilibrated structures, MD production runs of 20 ns (Systems 1–5) or 10 ns (System 6) were performed. All simulations were performed at 1 atm and 300 K, maintained with the Berendsen barostat and thermostat, respectively,[17] using periodic boundary conditions and the particle mesh Ewald method (grid spacing of 1Å) for treating long-range electrostatic interactions with a uniform neutralizing plasma. The SHAKE algorithm was used to keep bonds involving H atoms at their equilibrium length, allowing the use of a 2 fs time step for the integration of Newton's equations.

We applied steered MD to smoothly rotate the Arg565 side chain in the *ss*DAF-12 to force the contact with the DA. Rotation was performed by constant velocity multiple steered molecular dynamics (MSMD) simulation. In this study, the reaction coordinate λ was characterized by rotation of the CB-CG-CD-NE torsion angle, while the N-CA-CB-CG torsion angle was constrained to avoid undesired motions of the Arg565 side chain. The restrained coordinate was carried out with a constant force of 50 kcal mol^{-1} $degrees^{-1}$. On the other hand, rotation was performed using a constant force of 20 kcal mol^{-1} $degree^{-1}$ and a turning velocity of 0.1 degrees/ps. The initial structure for the MSMD was the final snapshot of the 10 ns MD simulation of the *ss*DAF-12/(25*R*)-Δ^7-DA complex. After the 2 ns of MSMD, the system was equilibrated as above and submitted to 10 ns of classical MD.

Analysis of results

The time evolution of the distances among the polar receptor residues and the polar ligand groups was monitored with the Ptraj module. Computed energetic contributions corresponded to the electrostatic energy (ele) and Van der Waals contributions (vdw) arising from bond, angle and dihedral terms in the force field, the sum of which gave the total gas phase binding energy (MM). In System 1 MM calculations were performed on 400 snapshots of the last 5 ns of the trajectory. In System 2, calculations were performed on the two LBMs observed, 900 snapshots of the 12–20 ns interval for LBM A and 900 snapshots of the 3–11 ns interval for LBM B. In System 3, calculations were performed on 330 snapshots of the 1.1–4.4 ns interval for LBM A, 630 snapshots of the 7.2–13.5 ns interval for LBM B and 220 snapshots of the 15–17.2 ns interval for LBM C. In Systems 4 and 5, calculations were performed on 400 snapshots of the last 5 ns of the trajectory. In System 6, calculations were performed on 130 snapshots of the 0.9–1.9 ns interval for LBM A, 690 snapshots of the 2–8.9 ns interval for LBM B and on 400 snapshots of the last 5 ns of the trajectory with the Arg565 side chain rotate (LBM C).

The thermodynamic integration method was performed to calculate the relative free energies of (25*R*)-Δ^7-DA and (25*S*)-Δ^7-DA. Conversion of the ligands to the 27-nor-Δ^7-DA was performed in solution and in the protein in three steps: first the electrical charge of the C-25 methyl group was removed; then the VDW values of C-25 methyl were deleted and replaced by the hydrogen value using soft potentials and last the charge of this hydrogen was introduced. We used 11 λ values (0.01, 0.1,

Table I
Thermodynamical and Structural Information of ce*DAF-12* and ss*DAF-12* Ligand Interactions

				Average distances (Å)[b]					MM energy (Kcal/mol)[c]		
System	Receptor	Ligand	LBM[a]	Gln638	Arg564	Arg598	Arg602	Thr613	Total	vdw	ele
1	ce*DAF-12*	(25R)-Δ^7-*DA*		2.92	4.36	4.49	4.62	–	−248.4	−57.0	−191.4
2	ce*DAF-12*	(25S)-Δ^7-*DA*	A	2.91	4.34	4.48	4.62	–	−268.1	−58.4	−209.7
			B	2.94	4.39	4.48	4.83	–	−250.0	−58.0	−192.0
3	ce*DAF-12*	27-nor-Δ^7-DA	A	2.92	4.33	4.39	4.40	–	−259.7	−57.0	−202.7
			B	2.92	4.36	4.36	4.42	–	−251.2	−55.9	−195.3
			C	2.93	4.31	4.29	4.43	–	−255.8	−56.8	−199.0
4	ce*DAF-12 Arg564Cys*	(25R)-Δ^7-*DA*		3.00	–	6.10	4.41	–	−152.5	−60.4	−92.1
5	ce*DAF-12 Arg598Met*	(25R)-Δ^7-*DA*		2.97	4.58	–	4.48	–	−191.6	−59.9	−131.7
6	ss*DAF-12*	(25R)-Δ^7-*DA*	A	2.88	5.26	3.99	–	3.72	−148.1	−63.3	−84.8
			B	2.96	5.64	3.98	–	3.77	−164.4	−63.7	−100.7
			C	2.99	3.89	4.02	–	3.95	−180.6	−63.7	−116.9

vdw, Van der Waals; ele, electrostatic.
[a]LBM, ligand binding mode.
[b]Average distances of Gln638 to the ligand oxygen atom of the C-3 ketone and of Arg564, Arg598, Arg602, and Thr613 to the carbon atom of the C-26 carboxylate.
[c]Interaction energy contributions computed using the MMGBSA method.

0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 y 0.99) and at each λ-value one simulation was carried out, consisting of 50 ps of equilibration followed by 500 ps of data collection over which the average $\Delta V/\Delta\lambda$ was calculated. Integrations over λ were carried out using the trapezoidal integration method.

RESULTS AND DISCUSSION

Construction of *ce*DAF-12/(25*R*)-Δ^7-DA model

The aminoacid sequence of the *ce*DAF-12 LBD and its alignment with the *ss*DAF-12 LBD using the Modeller program are shown in Figure 2c. Although the sequence similarity between them is 42%, the alignment is straightforward since several regions are highly conserved, such as helix H3. Furthermore, of the residues that form the LBP, 19 are conserved, five are replaced by similar aminoacids and only five present major modifications (Asn to Asp532, Leu to Phe544, Phe to Cys642, Trp to Phe610, and Val to Arg602). Thus, the *ce*DAF-12 LBP has two more polar residues than *ss*DAF-12, which are positioned around the ligand C-26 carboxylate group. Using this sequence alignment and the coordinates of the *ss*DAF-12/(25*R*)-Δ^7-DA crystal complex we constructed three *ce*DAF-12 models using Modeller. The models obtained resulted very similar, without significant differences on the positions of Asp532, Cys642, Phe610, Phe544 and Asp532 residues, but with a clear discrepancy with respect to the Arg602 side chain conformation [Supporting Information Fig. S1(a)]. While in certain models the Arg602 side chain points to the ligand cavity, taking thus part of the LBP, in others it is rotated away from the ligand position. In the first case, manual introduction of the (25*R*)-Δ^7-DA ligand in the LBP causes an overlap between the side chains of the residue and the ligand. Therefore, we used the Autodock 4.2 program to find the best combination of Arg602 and steroid side chain conformations. (25*R*)-Δ^7-DA was docked into the best ranked Modeller model considering rotatables the torsion angles of both Arg602 and of the steroid side chain. Using 200 runs of genetic algorithm and a rmsd equal to 1Å, 127 clusters were obtained. Eighteen of them had the steroid adequately oriented within the LBP and only seven had the Arg602 residue oriented toward the ligand cavity, taking part of the LBP. The best solution of the lowest energy cluster, and also the most probable, presented Arg602 accommodated under the steroid side chain with the C-26 carboxylate group far from the guanidinium moiety, but contacting the Arg598 and Arg564 residues [Supporting Information Fig. S1(b)]. On the other end, the C-3 keto group interacted with the Gln638. In this way, we obtained an initial structure of the *ce*DAF-12/(25*R*)-Δ^7-DA complex, in which the ligand occupied a similar position to that in the *ss*DAF-12 complex.

LBM analysis by MD

The LBMs in *ce*DAF-12 and *ss*DAF-12 receptors (systems 1–6, Table I) were studied by MD simulations. The selection of the corresponding initial structures is described in Material and Methods. All trajectories were analyzed by monitoring the time evolution and the average values of distances among ligand polar groups and all polar residues of the LBP. In the case of the C-3 keto group, the distance between its oxygen atom and the NE2 atom of Gln637 was selected; in the case of the C-26 carboxylate group, we calculated the distance between its carbon atoms and the CZ atom of arginine or OG1 atom of threonine residues. The dynamic behavior of the steroid side chain was inspected through the temporal

evolution of the d_1 (C20-C22-C23-C24), d_2 (C22-C23-C24-C25), and d_3 (C23-C24-C25-C26) torsion angles. While the d_1 and d_2 values depicted the global conformation of the steroid side chain, the d_3 values indicate the relative orientation of the C-26 carboxylate group. Moreover, the electrostatic (ele) and van der Waals (vdw) contributions to the total energy of the molecular mechanics (MM) force field were calculated in all cases. Although the MM energy does not include the solvation nor the entropic terms of the binding free energy, we considered that it was able to qualitatively estimate their relative ligand affinities as the receptor–ligand systems are quite similar, that is, even though the MM value itself may not have a physical meaning, comparison of values for the same receptor in different situations, is meaningful.

LBM in ceDAF-12/(25R)-Δ^7-DA complex

Visual inspection revealed a very stable *ce*DAF-12/(25*R*)-Δ^7-DA trajectory (System 1). We observed that after a rapid conformational change during the first stage of the simulation, a strong LBM with the three LBP arginines contacting the C-26 carboxylate group is reached (Fig. 3). The time evolution of C-26 carboxylate-arginine distances reflected this dynamical behavior and showed the stability of the LBM. The C-3 keto group forms a strong hydrogen bond with the Gln628 [Supporting Information Fig. S2(a)] and the C-26 carboxylate group interacts with Arg564, Arg598, and Arg602. Besides these ligand–receptor interactions, several protein electrostatic interactions are essential to stabilize the system. The Asp532 and the Arg598 residues form a very strong salt bridge interaction [Supporting Information Fig. S3(a)]. Moreover, the NE atom of the Arg602 forms a stable hydrogen bond with the oxygen atom of the backbone Arg598, keeping the first residue adequately oriented to the C-26 carboxylate ligand group. The Arg564 is also kept in this position by interaction with the oxygen atom of the backbone Met531. Regarding the steroid side chain, both d_1 and d_2 torsion angles show a narrow fluctuation around 180°, indicating that a fully extended conformation is maintained during the time scale of the simulation. Moreover, after the initial conformational change, the d_3 torsion angle is always around 300°, indicating that the C-26 carboxylate orientation is maintained. We also observed the presence of several water molecules close to the carboxylate ligand group and the arginine residues. Particularly, two water molecules form hydrogen bonds with the ligand carboxylate group that are established *ca.* 30% of the time (a H-bond was defined as present whenever the distance between the oxygen atoms involved in the interaction was less than 3.5Å). These water are exchanged with "external" water several times during the timescale of the MD simulation. Nevertheless, the presence of solvent molecules around this highly charged region of the LBD might play an important stabilizing role. Finally, the energetic analysis revealed that, although the ligand only has two polar groups, the electrostatic contribution results larger than the van der Waals contribution.

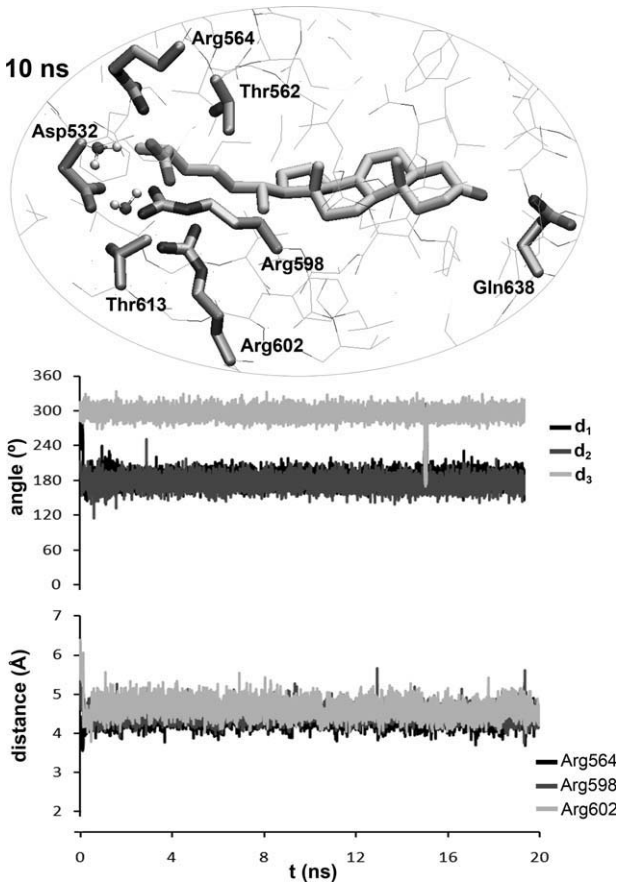

Figure 3

LBM analysis of the *ce*DAF-12/(25*R*)-Δ^7-DA complex (System 1). Lower panels: time evolution of distances between the ligand C-26 carboxylate and Arg564, Arg598, and Arg602 residues and time evolution of torsion angles of the steroid side chain. Upper panel: representative snapshot (10 ns) of the LBM of (25*R*)-Δ^7-DA.

Therefore, starting a MD analysis from the Autodock *ce*DAF-12/(25*R*)-Δ^7-DA structure, we found a very stable trajectory over 20 ns, in which a tight LBM was observed. The 10 ns snapshot protein coordinates of this trajectory were selected as a *ce*DAF-12 model and its secondary structure was analyzed with the DSSP method to ensure that it satisfied the elemental structural requisites.[18] Our homology model had 55% of its residues as part of α-helices and 2% as β-sheets, these values were very similar to those obtained with the *ss*DAF-12 crystal structure used as template (57% as α-helices and 2% as β-sheets). Moreover, the Ramachandran plot showed that only one residue was in the non-allowed region while the Verify3D evaluation[19] indicated that all residues had an adequate score.

Next, to further examine the *ce*DAF-12 model here obtained, we constructed other *ce*DAF-12 ligand systems

for which experimental activity data is available. First, the LBM of the more active (25*S*)-Δ^7-DA ligand was analyzed, and then the (25*R*)-Δ^7-DA binding mode in two *ce*DAF-12 mutants was evaluated.

LBM in the ceDAF-12/(25S)-Δ^7-DA complex

The (25*S*)-Δ^7-DA steroid was docked into the *ce*DAF-12 structure using Autodock. The best solution of the lowest and more probable cluster places the steroid in a similar position than the (25*R*)-Δ^7-DA ligand, with the C-3 keto group contacting the Gln368 and the side chain in the fully extended conformation (d_1 and d_2 equal to 180°), although with a different C-26 carboxylate orientation (d_3 equal to 60°). From this initial structure, 20 ns of MD simulation were performed and the LBM was analyzed (System 2). The visual inspection of the trajectory revealed that the (25*S*)-Δ^7-DA side chain is maintained in its original conformation (termed as LBM A) for the first 1 ns, when a significant conformational change takes place [Fig. 4(a)]. The d_1 angle values shift to *ca.* 60° and simultaneously the carboxylate group rotates in a way to continue opposing its oxygen atoms toward the Arg598 guanidinium group (d_3 around 60°). This (25*S*)-Δ^7-DA binding mode (termed as LBM B) remained stable for approximately 10 ns, after that the original conformation LBM A is again explored, reestablishing the fully extended steroid side chain. Interestingly, parallel to these conformational changes, the distance between the Arg602 and the C-26 carboxylate group was modified, being slightly lower in the LBM A than in the LBM B. The rest of the polar ligand–receptor interactions, and the exchange of several water molecules around the charged groups, do not exhibit significant differences with respect to System 1.

Thus, the MD analysis of the (25*S*)-Δ^7-DA binding from a *ce*DAF-12/(25*R*)-Δ^7-DA structure shows clearly two different LBMs which differ in the conformation of the steroid side chain. Remarkably, although in both LBM A and B the three LBP arginines interact with the carboxylate group, the ele contribution in the LBM A is 17.7 Kcal/mol larger than in the LBM B. In this way, we found that the conformation of the ligand with the side chain fully extended produces a more favorable stabilization than the folded conformation, and remarkably, this (25*S*)-Δ^7-DA binding mode produces a more favorable MM energy value than that observed for the (25*R*)-Δ^7-DA. It may be concluded that the configuration of the C-25 methyl affects the strength of the electrostatic ligand–receptor interactions, being more favorable in the *S* configuration. This finding may explain the relative activity of these isomeric DAs.

LBM in the ceDAF-12/27-nor-Δ^7-DA complex

To further evaluate the role of the C-25 methyl in the binding mode of DAs, we simulated the *ce*DAF-12 complex with the steroid lacking the C-25 methyl (27-nor-Δ^7-DA, System 3). The initial *ce*DAF-12/27-nor-Δ^7-DA structure was obtained by simple deletion of the C-25 methyl (replaced by hydrogen) in the 10 ns snapshot of the *ce*DAF-12/(25*R*)-Δ^7-DA trajectory. The 27-nor-Δ^7-DA binding mode resulted similar to those of the C-25 methyl compounds [Fig. 4(b)], with the three arginines interacting with the C-26 carboxylate ligand group during the time scale of the simulation. However, we found that the C-25 methyl absence does affect notably the dynamic behavior of the steroid side chain. In this case, three different conformations are explored along the trajectory: the fully extended (LBM A), the $d_1 = 60°$ torsioned (LBM B) and another d_1-torsioned but with values around 300° (LBM C), which was not observed previously. Therefore, this result suggests that the C-25 methyl could be involved in the reduction of the flexibility of the steroid side chain inside the LBP. Notably, all the LBM have similar ele contributions, indicating that in absence of the C-25 methyl the conformation of the steroid side chain does not influence the energetics of the polar interactions.

Finally, we applied the thermodynamic integration method to further evaluate the difference in binding free energy between both stereoisomers of the Δ^7 DA. The thermodynamic cycle construct is shown in Supporting Information Figure S4 confirming that no changes occurred in the LBM during 500 ps trajectories. The results show that both in water and in the protein environment, the C-25 methyl deletion implies a loss of energy in both (25*S*)-Δ^7-DA and (25*R*)-Δ^7-DA, although the difference between protein–water values results slightly larger in the S stereoisomers. In other words, the introduction of a methyl in both C-25 configurations increases the ligand affinity, but while for the *R* configuration this gain is only 0.9 kcal/mol, for the *S* configuration the values almost duplicate (1.7 Kcal/mol). Therefore, the binding energy of (25*S*)-Δ^7-DA is 0.8 kcal/mol stronger than the (25*R*)-Δ^7-DA binding energy, which is again consistent with the higher activity reported for the 25*S* stereoisomer. Although this relative free energy of binding is small, and of the order of magnitude of statistical errors associated with incomplete sampling observed usually in thermodynamic integration calculations,[20] since we have obtained converged trajectories to ligands which only differ in a methyl group, we expect that systematic errors are canceled and thus a correct prediction of DAs relative affinity can be made. On the other hand, other limitations of the thermodynamic integration method, such as those due to flaws in force fields (i.e., lack of polarization effects), are not expected to be significant since the ligands are very similar.

LBM in ceDAF-12/(25R)-Δ^7-DA mutants

As mentioned above, the substitution of Arg564 or Arg598 by non polar residues produces different effects

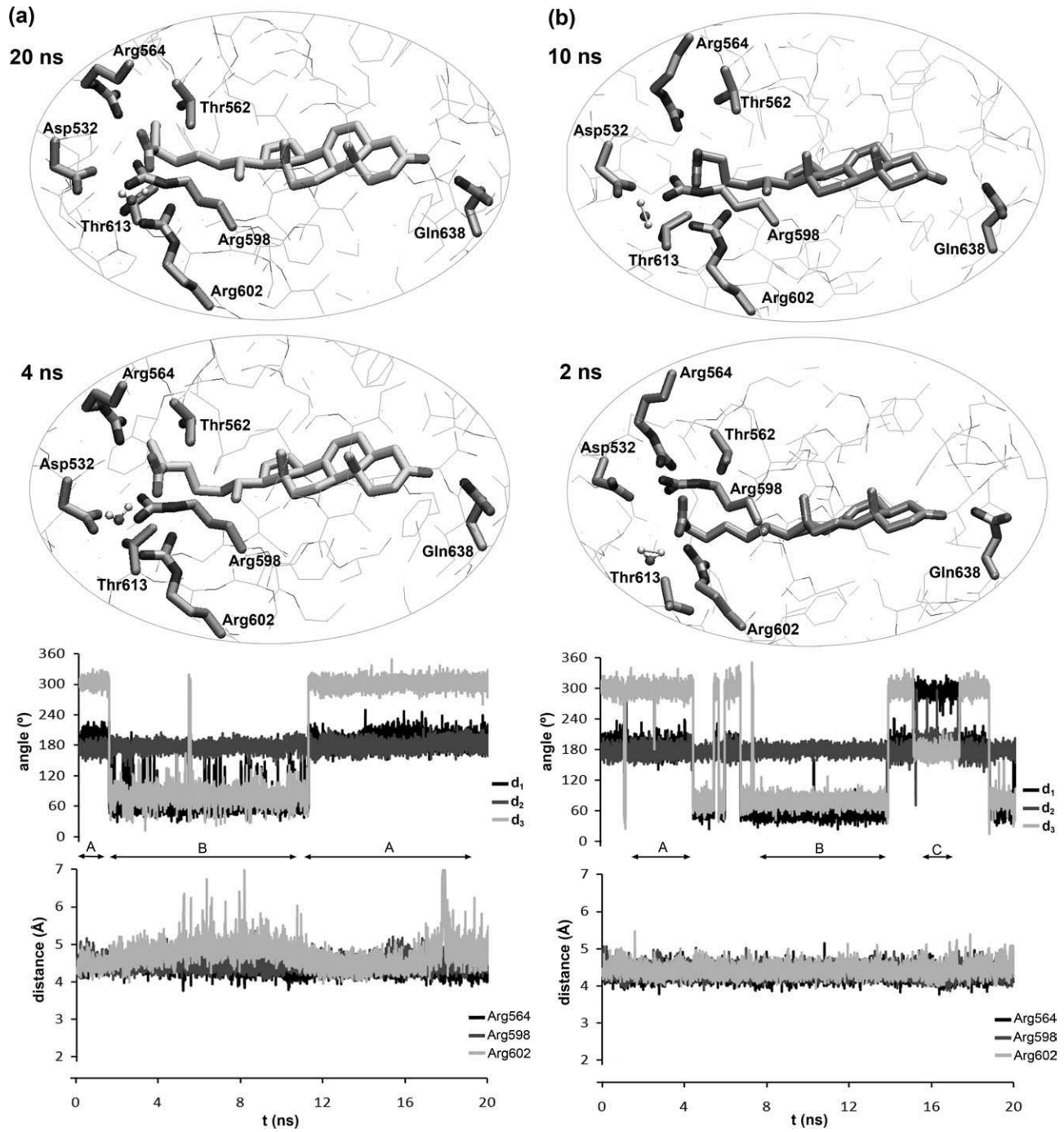

Figure 4

(**a**) LBM analysis of the *ce*DAF-12/(25*S*)-Δ^7-DA complex (System 2). Lower panels: time evolution of distances between the ligand C-26 carboxylate and Arg564, Arg598, and Arg602 residues and time evolution of torsion angles of the steroid side chain showing intervals for LBMs A and B. Upper panel: representative snapshots (20 and 4 ns) of the LBMs A and B of (25*S*)-Δ^7-DA. (**b**) LBM analysis of the *ce*DAF-12/27-nor-Δ^7-DA complex (System 3). Lower panels: time evolution of distances between the ligand C-26 carboxylate and Arg564, Arg598, and Arg602 residues and time evolution of torsion angles of the steroid side chain showing intervals for LBMs A, B and C. Upper panel: representative snapshots (2 ns and 10 ns) of the LBMs A and B of 27-nor-Δ^7-DA.

on the *ce*DAF-12/(25*R*)-Δ^7-DA transactivation activity. While the Arg564Cys mutant completely lost the (25*R*)-Δ^7-DA agonist effect, the Arg598Met mutation preserves over half of the wild type receptor activity. We used our *ce*DAF-12 model to study the effect of these mutations on the (25*R*)-Δ^7-DA binding mode. The mutant com-

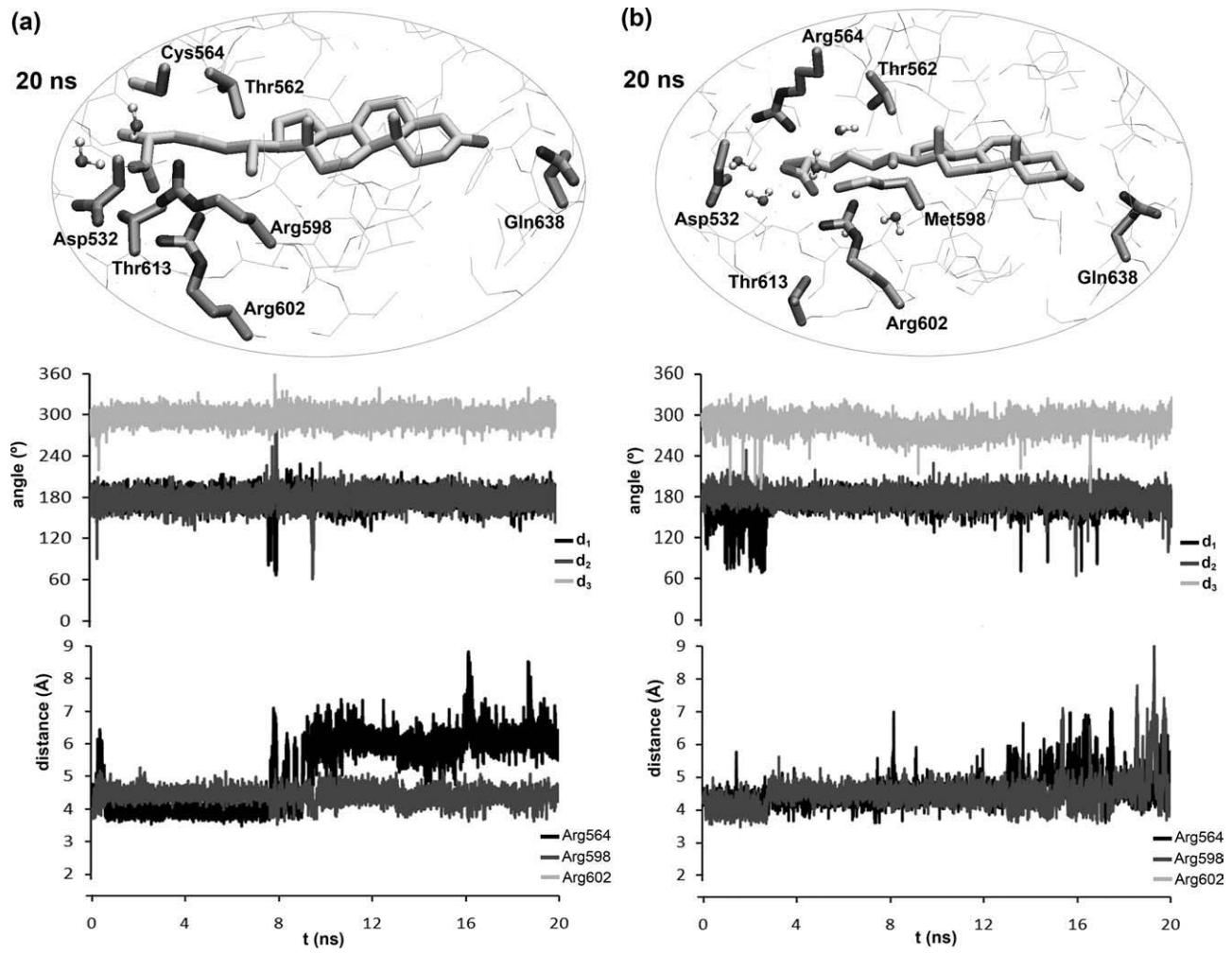

Figure 5

(**a**) LBM analysis of the Arg564Cys *ce*DAF-12/(25R)-Δ^7-DA complex (System 4). Lower panels: time evolution of distances between the ligand C-26 carboxylate and Arg564, Arg598, and Arg602 residues and time evolution of torsion angles of the steroid side chain. Upper panel: representative snapshot (20 ns) of the LBM of (25*R*)-Δ^7-DA. (**b**) LBM analysis of the Arg598Met *ce*DAF-12/(25*R*)-Δ^7-DA complex (System 5). Lower panels: time evolution of distances between the ligand C-26 carboxylate and Arg564, Arg598, and Arg602 residues and time evolution of torsion angles of the steroid side chain. Upper panel: representative snapshot (20 ns) of the LBM of (25*R*)-Δ^7-DA.

plexes were constructed from the 10 ns snapshot of *ce*DAF-12/(25*R*)-Δ^7-DA system by simple deletion of arginine side chains and introduction of cysteine and methionine side chains using the Tleap module of Amber.

The MD trajectory of the Arg564Cys mutant (System 4) shows a dramatic change of the overall LBM compared with the wild type receptor [Fig. 5(a)]. During the first 8 ns, the system conserved a conformation in which both arginines residues remained bound to the ligand, although it was not stable as the Arg598-ligand carboxylate interaction gradually disappeared reaching a binding mode that lacked a direct interaction between these polar groups. Moreover, alterations in the Arg533-Asp598 distance reflected a less tight binding mode [Supporting Information Fig. S3(d)]. As a consequence, the ele contribution of the Arg564Cys mutant is very small compared with the other *ce*DAF/12 systems. Since the Arg–Cys substitution produces a larger cavity around the ligand carboxylate, many water molecules access the carboxylate ligand position, perturbing the interaction with the Arg598 and affecting the ligand binding. In this way, our *ce*DAf-12 model suggests that the lack of transactivation activity of (25*R*)-Δ^7-DA in the *ce*DAF-12 Arg598Cys mutant could reside in the inability of the receptor to adequately recognize the ligand molecule.

Instead, in the case of Arg598Met mutant (System 5), we found smaller effects on the LBM obtained. A rapid alternation between the fully extended and the d_1-torsioned conformations of the steroid side chain was observed during the first 4 ns [Fig. 5(b)]. Then, a smaller but abrupt change in the Arg598-ligand carboxylate

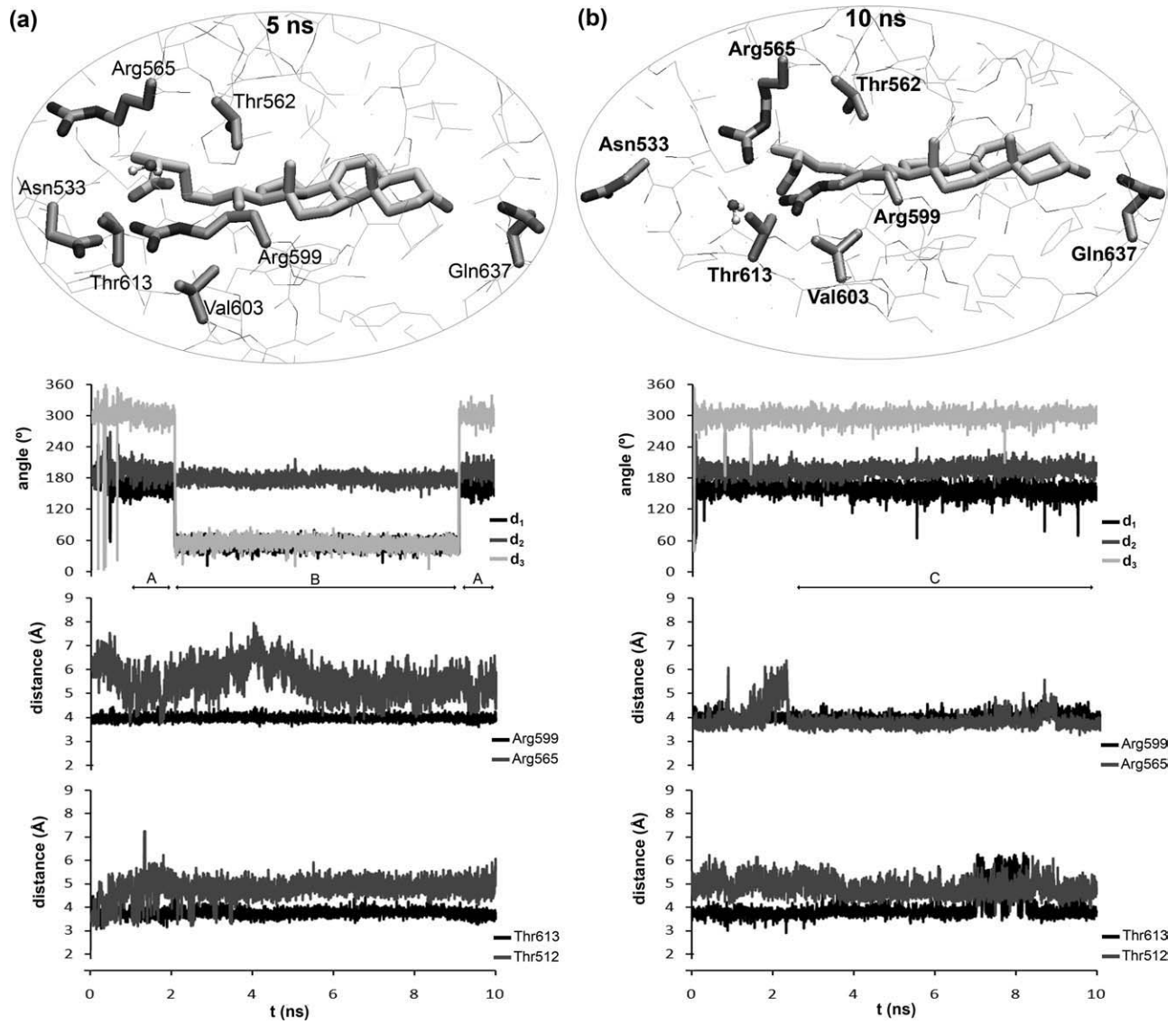

Figure 6

(**a**) LBM analysis of the *ss*DAF-12/(25R)-Δ^7-DA complex. Lower panels: time evolution of distances between the ligand C-26 carboxylate and Arg565, Arg599, Thr512, and Thr613 and time evolution of torsion angles of the steroid side chain. Upper panel: representative snapshot (5 ns) of the LBM of (25*R*)-Δ^7-DA. (**b**) LBM analysis of the *ss*DAF-12/(25R)-Δ^7-DA complex after steered MD to smoothly rotate the Arg565 side chain approaching the guanydinium group to the C-26 carboxylate position. Lower panels: time evolution of distances between the ligand C-26 carboxylate and Arg565, Arg599, Thr512, and Thr613 and time evolution of torsion angles of the steroid side chain. Upper panel: representative snapshot (1 ns) of the LBM of (25*R*)-Δ^7-DA.

occurred and the system evolved to a conformation in which the steroid side chain remains fully extended and the arginines-ligand carboxylate distances gradually tend to increase. Although there are some events in which the interaction between the arginines and the ligand carboxylate group are actually lost close to the end of the simulation, the system results much more stable and the ligand more strongly bound compared with System 4. In this way, the ele contribution of the Arg598Met mutant (System 5) results larger than in the Arg564Cys mutant (System 4), but not as large as in the wild type receptor (System 1), in agreement with the effects observed for these *ce*DAF-12 mutations on the transactivation activity.

LBM in the ssDAF-12/(25R)-Δ^7-DA complex

Finally, to validate our model and for comparative purposes, we studied the LBM in the *ss*DAF-12/(25*R*)-Δ^7-DA complex (System 6) by MD, and compared the results with the crystal data and with the results obtained for the *ce*DAF-12 receptor. First, the results shows that the use of the *ss*DAF-12/(25*R*)-Δ^7-DA crystal structure

(pdb:3GYU) as initial structure generates a trajectory that does not reproduce the LBM observed in the crystal structure (see Methods). Particularly, the interaction between the Asn533 and the Arg599 is never established, allowing the entrance of several water molecules into the LBP and perturbing the ligand–receptor interactions (data not shown). Instead, when the MD was started from the crystal structure with the other Asn553 rotamer, we obtained a stable trajectory in which the overall LBP conformation was conserved during the MD simulation of 10 ns very close to the initial structure. The interaction between Arg599 and Asn533 is always present [Supporting Information Fig. S3(f)] and stable interactions between the ligand carboxylate group and both Arg599 and Thr613 are formed [Fig. 6(a)]. Moreover, the 3-keto group forms a strong hydrogen bond with the Gln637 residue during the whole time scale of the simulation [Supporting Information Fig. S2(f)]. Like in the crystal structure, the Arg565 residue remains bound to the H1–H2 loop backbone, without interacting with the ligand. The MD simulation also shows that the steroid side chain alternates between a fully extended conformation (LBM A), found in the beginning and at the end, and a $d_1 = 60^\circ$ torsioned conformation (LBM B) in the rest of the trajectory. Remarkably, in this system the torsioned conformation has a larger ele contribution.

Next, taking into account that in the simulated *ce*DAF-12 systems the Arg565 and the ligand C-26 carboxylate group actually interact, we were interested on analyzing the ability of this residue to change its side chain orientation to contact the ligand group in the *ss*DAF-12. With this purpose, we took the last snapshot of the above *ss*DAF-12 trajectory and applied steered MD to smoothly rotate the Arg565 side chain approaching the guanidinium group to the C-26 carboxylate position. Then, the restrain conditions were removed and 10 ns of classical MD simulation were performed. The results obtained show that the Arg565 actually is able to form a strong interaction with the ligand carboxylate, without affecting the overall LBP structure and the other ligand–receptor interactions (LBM C) [Fig. 6(b)]. The only remarkable change resides in the loss of the interaction between the Asn533 and the Arg599 [Supporting Information Fig. S3(g)], although this does not affect the Arg599 interaction with the ligand. Moreover, the additional arginine–ligand interaction largely increased the ele contribution of this LBM, giving a total MM energy more favorable than that obtained in LBM A and B.

In summary, using MD simulations we have reached different stable binding modes of (25*R*)-Δ^7-DA in the *ss*DAF-12 receptor: two in which the ligand interacts only with Arg599, as in the crystal structure (LBM A and B), and another with the Arg565 rotated and forming also a polar interaction with the carboxylate group (LBM C), which clearly results in a more energetically favorable ligand–receptor recognition. By comparison of the MM energy values of *ce*DAF-12 and *ss*DAF-12 systems, we conclude that the additional Arg602-ligand interaction contributes largely to the binding. Thus, although the vdw contribution is slightly smaller in the *ce*DAF-12 complex, probably due to the loss of a non polar valine in the LBP, the total MM energy is more favorable than in the *ss*DAF-12 complex. This result is consistent with the experimental results obtained by Wang *et al.* on the relative activity between these receptors.[7]

CONCLUSIONS

The *ce*DAF-12 initial structure obtained from the *ss*DAF-12/(25*R*)-Δ^7-DA crystal structure allowed us to simulate its complexes with two ligands for which the transactivation activities have been reported and the model was able to explain the experimental results at the molecular level. Thus, the higher activity of (25*S*)-Δ^7-DA compared with the (25*R*)-Δ^7-DA isomer could be related to the strong ligand–receptor electrostatic interaction produced in the fully extended conformation of the (25*S*)-isomer side chain. Moreover, since this side chain results more flexible within the LBP, the entropic penalty due to binding should be lower than in the more rigid (25*R*) isomer side chain, increasing its affinity. Simulation of key *ce*DAF-12 mutants bound to the (25*R*)-Δ^7-DA, led us to propose that the lack of activity of the *ce*DAF-12 Arg564Cys is a consequence of its inability to achieve an adequate ligand recognition due to the presence of several water molecules in the LBP that interfere in the polar interaction between the protein and the C-26 carboxylate group. In contrast, the intermediate transactivation activity of the Arg598Met mutant could be explained by the less energetically favorable, but stable, LBM observed in this system. Taken together, the results obtained here suggest that our *ce*DAF-12 model may be used reliably as an initial structure in molecular modeling studies. In this sense, we have used it to investigate the LBM of an analog of the Δ^7-DA with a simplified steroid side chain (27-nor-DA). The model showed that the lack of the C-25 methyl group increases the flexibility of the side chain, but this does not impede the simultaneous strong interaction of the carboxylate group with the three arginines, predicting that this steroid is a putative *ce*DAF-12 ligand.

Since many biological processes are conserved between humans and *C. elegans*, the active participation of *ce*DAF-12 in the nematode life cycle, regulating both dauer formation and adult longevity, converts this NR in a valuable molecular target for the study of elemental process, such as development and adult longevity. We believe that until a crystal structure is available, the molecular model of *ce*DAF-12 will allow not only the design and discovery of new ligands, but also a more detailed study of the receptor basis of action, LBMs and other aspects of the highly complex NR activity, such as homodimerization and cofactors recruitment events. It should

be stressed that the discovery of antagonist ligands for this receptor will have a major impact on the field.

ACKNOWLEDGMENTS

We thank Dr. Marcelo Martí (INQUIMAE) for helpful discussions. LDA and PAM thank CONICET (Argentina) for a postdoctoral fellowship.

REFERENCES

1. Kaletta T, Hengartner MO. Finding function in novel targets: *C. elegans* as a model organism. Nat Rev Drug Discov 2006;5:387–398.
2. Motola DL, Cummins CL, Rottiers V, Sharma KK, Li T, Li Y, Suino-Powell K, Xu HE, Auchus RJ, Antebi A, Mangelsdorf DJ. Identification of ligands for DAF-12 that govern dauer formation and reproduction in *C. elegans*. Cell 2006;124:1209–1223.
3. Beckstead RB, Thummel CS. Indicted: worms caught using steroids. Cell 2006;124:1137–1140.
4. Hochbaum D, Zhang Y, Stuckenholz C, Labhart P, Alexiadis V, Martin R, Knölker HJ, Fisher AL. DAF-12 regulates a connected network of genes to ensure robust developmental decisions. PLoS Genet 2011;7:e1002179.
5. Antebi A, Yeh WH, Tait D, Hedgecock EM, Riddle DL. daf-12 encodes a nuclear receptor that regulates the dauer diapause and developmental age in *C. elegans*. Genes Dev 2000;14:1512–1527.
6. Gerisch B, Weitzel C, Kober-Eisermann C, Rottiers V, Antebi A. A hormonal signaling pathway influencing *C. elegans* metabolism, reproductive development, and life span. Dev Cell 2001;1:841–851.
7. Wang Z, Zhou XE, Motola DL, Gao X, Suino-Powell K, Conneely A, Ogata C, Sharma KK, Auchus RJ, Lok JB, Hawdon JM, Kliewer SA, Xu HE, Mangelsdorf DJ. Identification of the nuclear receptor DAF-12 as a therapeutic target in parasitic nematodes. Proc Natl Acad Sci USA 2009;106:9138–9143.
8. Martin R, Entchev EV, Däbritz F, Kurzchalia TV, Knölker HJ. Synthesis and hormonal activity of the (25*S*)-cholesten-26-oic acids—potent ligands for the DAF-12 receptor in *Caenorhabditis elegans*. Eur J Org Chem 2009:3703–3714.
9. Sharma KK, Wang Z, Motola DL, Cummins CL, Mangelsdorf DJ, Auchus RJ. Synthesis and activity of dafachronic acid ligands for the *C. elegans* DAF-12 nuclear hormone receptor. Mol Endocrinol 2009;23:640–648.
10. Martin R, Entchev EV, Kurzchalia TV, Knölker HJ. Steroid hormones controlling the life cycle of the nematode *Caenorhabditis elegans*: stereoselective synthesis and biology. Org Biomol Chem 2010;8:739–750.
11. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
12. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comp Chem 1998;19:1639–1662.
13. Case DA, Darden TA, Cheatham III TE, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts BP, Wang B, Hayik S, Roitberg A, Seabra G, Kolossvai I, Wong KF, Paesani F, Vanicek J, Liu J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Cai Q, Ye X, Wang J, Hsieh M-J, Cui G, Roe DR, Mathews DH, Seetin MG, Sagui C, Babin V, Luchko T, Gusarov S, Kovalenko A, Kollman PA. Amber, Version 11. San Francisco: University of California; 2010.
14. Weichenberger CX, Sippl MJ. NQ-Flipper: recognition and correction of erroneous asparagine and glutamine side-chain rotamers in protein structures. Nucleic Acids Res 2007;35(Web Server issue): W403–W406.
15. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA, Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA. Gaussian, Version 03, Revision C. Wallingford: Gaussian Inc.; 2004.
16. Cheatham TE, Cieplak P, Kollman PA. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. Biomol Struct Dyn 1999;16:845–862.
17. Berendsen HJC, Postma JPM, Van Gunsteren WF, DiNola A, Haak JR, Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
18. Kabsch W, Sander C, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 1983;22:2577–2637.
19. Luthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature 1992;356:83–85
20. Michel J, Foloppe N, Essex JW. Rigorous free energy calculations in structure-based drug design. Mol Inform 2010;29:570–578