Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry Letters 13 (2003) 1865–1868

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Binding Mode of 6ECDCA, a Potent Bile Acid Agonist of the Farnesoid X Receptor (FXR)

Gabriele Costantino, Antonio Macchiarulo, Antonio Entrena-Guadix,[†] Emidio Camaioni and Roberto Pellicciari*

Dipartimento di Chimica e Tecnologia del Farmaco, Università di Perugia, Via del Liceo 1, 06123 Perugia, Italy

Received 28 November 2002; revised 18 January 2003; accepted 13 March 2003

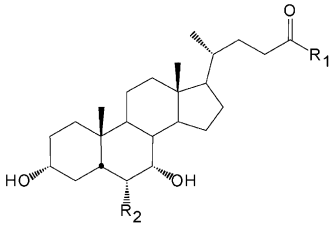
Abstract—Based on the folding conservation across the nuclear receptor superfamily and the sequence homology with RAR-γ, we report the construction of a three dimensional model of the ligand binding domain of FXR. The model is exploited for the elucidation of the binding mode of 6α-ethyl-chenodeoxycholic acid. The results of the docking experiments give quite clear indications that the bile acid derivative would bind the receptor in a mode significantly different than that observed for agonists of other nuclear receptor superfamily.

Introduction

We have recently reported the synthesis and the preliminary evaluation of a series of 6α-alkyl-substituted analogues of chenodeoxycholic acid (**1**, CDCA; Fig. 1),[1] among which 6α-ethyl-chenodeoxycholic acid (**3**, 6-ECDCA) turned out to be the most potent steroid agonist so far reported for the Farnesoid X Receptor (FXR). FXR belongs to the nuclear receptor superfamily of intracellular ligand-activated transcription factors that are involved in many physiological, developmental and toxicological processes.[2] Members of this superfamily are constituted by three functional conserved domains: an N-terminus dimerization domain, a central DNA binding domain, and a C-terminus ligand binding domain (LBD).[3] On the basis of sequence alignment analysis[4] and crystallization studies,[5] it has been proposed an highly conservative fold of the ligand binding domain (LBD).

A number of evidences have indicated a role of FXR in regulating the homeostatic control of lipid metabolism. In particular, FXR constitutes a feed-back regulator of the hepatic biosynthesis of bile acids from cholesterol.[6] This regulation occurs through the suppression of the cytochrome P450 gene encoding cholesterol 7α-hydroxylase (CYP7A1) which is the rate-limiting enzyme of the bile acids biosynthetic pathway, the induction of the bile salt export pump (Bsep) gene encoding the pump for the active transport of bile acids across the membrane of hepatocytes, and the induction of the intestinal bile acid binding protein (IBABP).

FXR also regulates the expression of several genes involved in lipoprotein metabolism. Indeed, FXR is a



Code	R_1	R_2	$EC_{50}(\mu M)$
1	OH	H	8.66 ± 0.45 [a]
2	OH	CH_3	0.75 ± 0.08 [a]
3	OH	CH_2CH_3	0.099 ± 0.01 [a]
4	OH	$CH_2CH_2CH_3$	1.11 ± 0.13 [a]
5	OH	Benzyl	> 30 [a]
6	$NHCH_2CO_2H$	H	3 [b]
7	$NHCH_2CH_2SO_3H$	H	3 [b]

Figure 1. Structures and activities of 6α-alkyl-chenodeoxycholic acids.
[a] see ref 1.
[b] see ref 6b.

*Corresponding author. Fax: +39-075-585-5124; e-mail: rp@unipg.it
[†]On leave from Departamento de Quimica Farmaceutica y Organica. Universidad de Granada (Spain).

0960-894X/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0960-894X(03)00281-6

```
                                                  *   *
LBD/RAR : LSPQLEELITKVSKAHQETFPSLCQLGKYTTNSSADHRVQLDLGLWDKFSELATKCIIKIVEFAKRLPGFTGLSIADQITLLKAACLDILM
LBD/FXR : -KTELTPDQQTLLHFIMDSYNK-QRMPQEITNKILKEEFSAE-ENFLILTEMATNHVQVLVEFTKKLPGFQTLDHEDQIALLKGSAVEAMF

                                          +
LBD/RAR : LRICTRYTPEQDTMTFSDGLTLNRTQMHNAGFGP--LTDLVFAFAGQLLPLEMDDTETGLLSAICLICGDRMDLEEPEKVDKLQEPLLEAL
LBD/FXR : LRSAEIFN---KKLPSGHSDLLE--ERIRNSGISDEYITPMFSFYKSIGELKMTQEEYALLTAIVILSPDRQYIKDREAVEKLQEPLLDVL

                                              +   *         *       +
LBD/RAR : RLYARRRRPSQPYMFPRMLMKITDLRGISTKGAERAITLKMEIPGPMPPLIREMLENP
LBD/FXR : QKLCKIHQPENPQHFACLLGRLTELRTFNHHHAEMLMSWRVND-HKFTPLLCEIWDVQ
```

Figure 2. Alignment of the figure of the LBDs of RAR-γ and FXR. Identical and similar aligned residues are gray shaded. Residues poorly scored in Verify3D are marked with a line. Residues defining the hydrophobic pocket are marked by a star. Residues involved in hydrogen bonding interactions are marked by a plus sign.

positive modulator of apoC-II[7] and the phospholipids transfer protein (PLTP) genes,[8] but a negative regulator of apoA-I gene.[9]

As part of our research project, we were interested in gaining insight into the binding mode of bile acid derivatives to the LBD of FXR. Indeed, the understanding of the molecular basis of such interactions would significantly aid the design of novel and potent modulators of FXR and help to improve the knowledge of the molecular basis of agonist-induced FXR activation mechanism.

A 3-D model of FXR was constructed on the basis of the sequence homology relationship with the LBD of the retinoic acid receptor (RAR-γ; pdb 2ldb).[5a]

The sequences of the ligand binding domains of human FXR and of human RAR-γ (pdb code: 2lbd) share 23% of residue identities and 44% of sequence similarity (Fig. 2).

Where missing, secondary structures were predicted using the PHD server. The alignment was carefully checked to avoid gap insertion where conserved secondary structure motifs were present. Insight-II software was used to build a 3-D model of the LBD of human FXR. The model thus obtained was submitted to a minimization protocol using the Charmm22 forcefield as implemented in Insight-II. During the minimization, the G-Born approximation was used to simulate the solvent environment. The energy minimization was performed using a cycle of 500 steps of Steepest Descent algorithm followed by several cycles of Conjugate Gradient algorithm until a gradient of 0.05 *K*cal/mol•Å^2 was reached. A geometric validation of the structure was carried out by using the Procheck server.[11] Where present, bad geometries were manually corrected and the structure minimized again with the above protocol. The final model was solvated within a 40Å radius sphere of water molecules starting from its centre of mass using the soak protocol of Insight-II. A total of 7.505 water molecules were added. The solvated model was imported into NAMD v.2.5b1 program for molecular dynamic calculations.[12]

Briefly, 1000 steps of Conjugate Gradient were performed applying fixed constrains on the atomic coordinates of solute. This optimization was followed by 100ps of dynamic simulation at 300 K constraining the solute. Finally, a molecular dynamic of 300 ps at 300 K was performed on the entire system without constrains. In all NAMD calculations, spherical boundary conditions were employed in order to keep the water near the protein. The last frame of the molecular dynamic simulation was stored and minimized with 1000 steps of Conjugate Gradient. The reliability of the obtained protein folding was assessed by using the Verify3D algorithm[13] and the Ramachandran plot. The model has more than 98% of residues in allowed regions of the Ramachandran plots. A short loop of the LBD of FXR (Pro270-Ala280) falls in the negative region of the Verify3D plot (Fig. 2). However, this loop corresponds to a poorly scored loop of the LBD of RAR-γ, which is not involved in the definition of the binding site of the receptor (Fig. 3).

Since we used the agonist-bound conformation of the LBD of RAR-γ as template, we expected a starting structure of the LBD of FXR near its agonist-bound conformation. The binding site of bile acids, identified

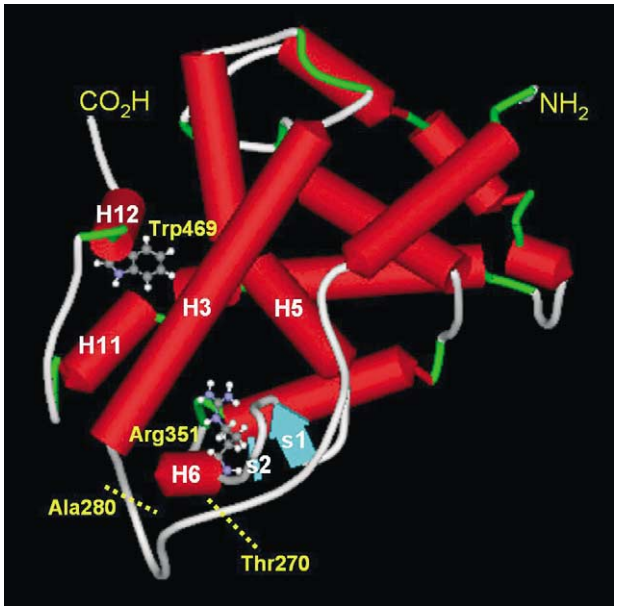


Figure 3. 3-D Model of the LBD of human FXR.

on the basis of the homology with known nuclear receptors, appears as a pocket of 687 Å^3 composed by residues belonging to the β-hairpin s1-s2 and helices H3, H5, H6, H11 and H12 on which the activation factor-2 (AF-2) is located (Fig. 3).

Docking experiments of 6-ECDCA (**3**) into the LBD of FXR were performed using Autodock v.3.0.[14] Briefly, 100 runs were carried out using the Lamarckian genetic algorithm with a population size of 100 individuals, 27,000 generations and 2,500,000 energy evaluations. The search was conducted in a grid of 42 points per dimension and a step size of 0.375 centered on the coordinates of the centre of mass of the ligand binding pocket of FXR.

Compound **3** was built using the sketch module of Cerius-2[10] and minimized using the Universal force-field v.1.2[15] with the Smart Minimizer protocol of the Open Force Field module (OFF). Atomic charges were calculated using the semi-empirical Mopac/AM1 method.

Docking results were clustered and ranked in terms of binding energy and occurrence for a given docked conformation. The analysis of the energetically most favorable solution reveals a binding disposition of 6-ECDCA (**3**) similar to that adopted by the trans-retinoic acid into the binding site of RAR-γ (pdb code: 2lbd). Indeed, **3** binds the LBD of FXR by disposing the acidic head towards the core of the receptor, by making salt bridge interaction with an arginine residue (Fig. 4).

The ring A of **3** points towards helix H12 and its 3α-hydroxyl group is inserted in a pocket formed by Trp469, His 447 and Phe461, in such a way of eventually stabilizing the functional active agonist-bound conformation of the LBD of FXR. Remarkably, Trp469 is conserved with Trp443 on the AF-2 sequence of LXRα which has been previously demonstrated, by mutagenesis experiments, to be involved in the binding of oxysterols.[16]

The 6α-ethyl moiety lays into an hydrophobic cage constituted by Phe284, Thr288, Leu451 and Phe461. It is worth noting that Phe284, conserved on the helix 3 of LXRs and PPARs sequences, has been reported as a key residue in ligand binding of human LXR-α.[17]

We have also analyzed the effect of FXR mutations across species[18] on the binding of 6-ECDCA (**3**). Indeed the double mutation Asn354/Lys366 and Ile372/Val384 has been reported as the origin of the higher affinity of CDCA to human FXR over the murine FXR. Since CDCA (**1**) and 6-ECDCA (**3**) share a similar chemical scaffold, a similar behavior is expected also for compound **3** in the affinity to the murine FXR. The Ile372/Val384 mutation is located far from the binding pocked of FXR thus indicating a possible indirect mechanism in affecting the ligand binding. Asn354 is directly involved in stabilizing the active side chain conformation of Arg351 for ligand binding. The replacement of Asn354 with Lysine leads to the insertion of a positive charge near the equally positively charged Arg351. As a consequence, Arg351 is likely to modify the conformation of its side chain into an inactive one.

The volume of the FXR binding pocket and 6-ECDCA (**3**), as defined by their molecular surfaces, are 687 Å^3 and 384 Å^3, respectively. Structure activity relationships have highlighted the existence of a small hydrophobic pocket in the receptor accommodating small alkyl groups in position 6 of the steroidal ring of CDCA [Ethyl (**3**) > *n*-Propyl (**4**) > Benzyl (**5**)].[1]

By means of volume comparison, carried out using the Surfnet program,[19] it is possible to observe that the 6-ethyl derivative (**3**) forms six crevices between its molecular shape and the binding cavity of FXR (Fig. 5).

In particular, the largest crevice (163 Å^3) is located near the carboxylic group while the smallest pocket (7 Å^3) is placed near the ethyl moiety of 6-ECDCA. This latter pocket could hardly fit substituents larger than ethyl (**3**) such as n-propyl (**4**) and benzyl (**5**) ones. On the other hand, the largest pocket (163 Å^3) near the carboxylic group could easily guest the glycine (42 Å^3) and taurine

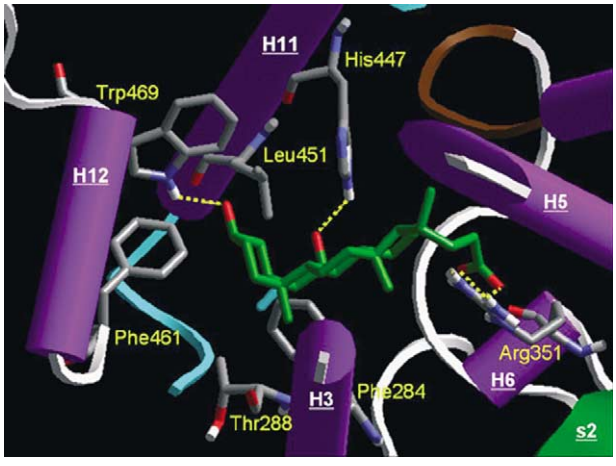


Figure 4. Docking result of 6-ECDCA.

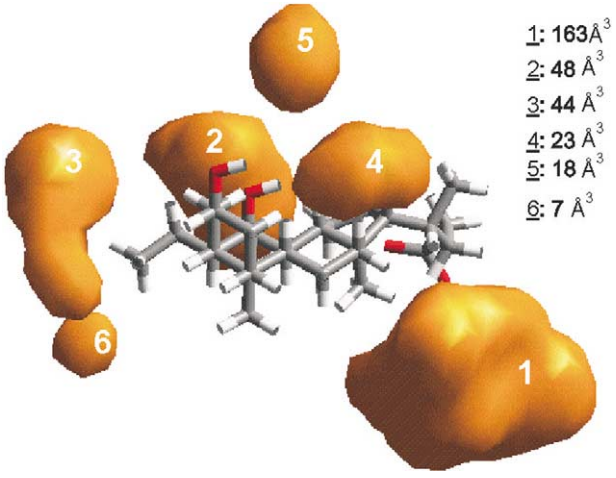


Figure 5. Volumes and location of the six pockets formed by 6-ECDCA in the binding site of FXR.

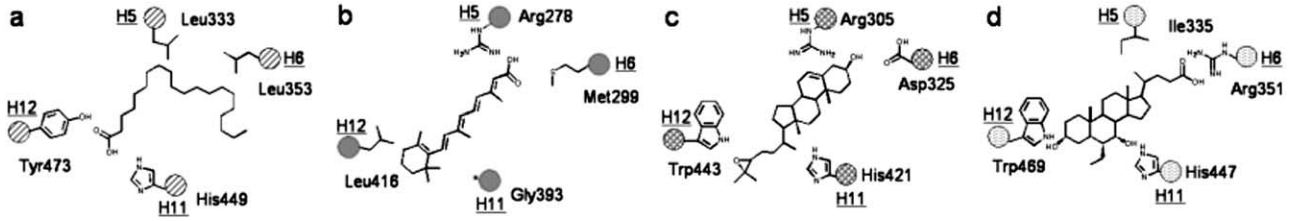


Figure 6. Schematic binding mode of different nuclear receptor agonists. (a) EPA. (b) T-RA. (c) Oxysterols. (d) Bile acids.

(69 Å^3) moieties of the glycol- (**6**) and tauro- (**7**) conjugates of CDCA.

Indeed, it has been reported that the glycine (**6**) and taurine (**7**) conjugates of CDCA activate FXR in tissue that express bile acid transporters (IBAT).[6b]

The present results can be commented on the basis of available data on the binding mode of nuclear receptor agonists, in particular trans-retinoic acid (T-RA) and fatty acids (eicosapentaenoic acid, EPA), to nuclear receptors. X-ray structures of such complexes are available respectively for RAR-γ[4a] and PPAR-γ.[4d]

Although agonists for these receptors are not steroidal, some conclusions can be drawn for comparison purposes. In the case of PPAR-γ agonist, the distal acidic moiety of EPA is found to interact with the Tyr473 residue located in the AF-2 helix (Fig. 6a).

In contrast, T-RA binds the RAR-γ receptor by disposing the acidic head towards the core of the receptor, thus making salt bridge interaction with Arg278 (Fig. 6b). Furthermore, a recent combined homology modeling-docking study[16] identified a putative binding mode for the steroidal oxysterol ligands to LXR-α receptor. Although oxysterols are not acidic agonists, their polar side chain directly made interaction with the AF-2 helix (Fig. 6c), in a similar way to that of PPAR agonists. Our present findings delineate a further different binding mode for the steroid 6-ECDCA (**3**) ligand, which seems to prefer a retinoic-like disposition, where the acidic side chain is instrumental to make strong salt-bridge interaction with Arg351 (Fig. 6d). Thus, these combined data indicate that while the binding pocket topology is conserved across the nuclear receptor superfamily, there could be multiple mechanisms by which ligands interact with the binding pocket to activate the receptor.

These results will be used to address the design and synthesis of new potent and selective FXR modulators endowed by anticholestatic activity.

Acknowledgements

Financial support from GSK (North Carolina, USA) is gratefully acknowledged. A. Entrena-Guadix gratefully acknowledges the 'Secretaria de Estado de Universidades y Educacion' of Spain for financial support.

References and Notes

1. Pellicciari, R.; Fiorucci, S.; Camaioni, E.; Clerici, C.; Costantino, G.; Maloney, P. R.; Morelli, A.; Parks, D. J.; Willson, T. M. *J. Med. Chem.* **2002**, *45*, 3569.
2. Laudet, V. *J. Mol. Endocrinol.* **1997**, *19*, 207.
3. Evans, R. M. *Science* **1988**, *240*, 889.
4. Wurtz, J. M.; Bourguet, W.; Renaud, J. P.; Vivat, V.; Chambon, P.; Moras, D.; Gronemeyer, H. *Nat. Struct. Biol.* **1996**, *3*, 206.
5. (a) Renaud, J. P.; Rochel, N.; Ruff, M.; Vivat, V.; Chambon, P.; Gronemeyer, H.; Moras, D. *Nature* **1995**, *378*, 681. (b) Williams, S. P.; Sigler, P. B. *Nature* **1998**, *393*, 392. (c) Nolte, R. T.; Wisely, G. B.; Westin, S.; Cobb, J. E.; Lambert, M. H.; Kurokawa, R.; Rosenfeld, M. G.; Willson, T. M.; Glass, C. K.; Milburn, M. V. *Nature* **1998**, *395*, 137. (d) Xu, H. E.; Lambert, M. H.; Montana, V. G.; Parks, D. J.; Blanchard, S. G.; Brown, P. J.; Sternbach, D. D.; Lehmann, J. M.; Wisely, G. B.; Willson, T. M.; Kliewer, S. A.; Milburn, M. V. *Mol. Cell* **1999**, *3*, 397.
6. (a) 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. *Science* **1999**, *284*, 1362. (b) Parks, 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. *Science* **1999**, *284*, 1365. (c) Wang, H.; Chen, J.; Hollister, K.; Sowers, L. C.; Forman, B. M. *Mol. Cell* **1999**, *3*, 543.
7. Kast, H. R.; Nguyen, C. M.; Sinal, C. J.; Jones, S. A.; Laffitte, B. A.; Reue, K.; Gonzalez, F. J.; Willson, T. M.; Edwards, P. A. *Mol. Endocrinol.* **2001**, *15*, 1720.
8. Urizar, N. L.; Dowhan, D. H.; Moore, D. D. *J. Biol. Chem.* **2000**, *275*, 39313.
9. Claudel, T.; Sturm, E.; Duez, H.; Torra, I. P.; Sirvent, A.; Kosykh, V.; Fruchart, J. C.; Dallongeville, J.; Hum, D. W.; Kuipers, F.; Staels, B. *J. Clin. Invest.* **2002**, *109*, 961.
10. Cerius-2 and Insight-II, Accelrys, San Diego, CA.
11. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283.
12. Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comp. Physic.* **1999**, *151*, 283.
13. Luthy, R.; Bowie, J. U.; Eisenberg, D. *Nature* **1992**, *356*, 83.
14. Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comp. Chem.* **1998**, *19*, 1639.
15. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.
16. Spencer, T. A.; Li, D.; Russel, J. S.; Collins, J. L.; Bledsoe, R. K.; Consler, T. G.; Moore, L. B.; Galardi, C. M.; McKee, D. D.; Moore, J. T.; Watson, M. A.; Parks, D. J.; Lambert, M. H.; Willson, T. M. *J. Med. Chem.* **2001**, *44*, 886.
17. Urban, F., Jr.; Cavazos, G.; Dunbar, J.; Tan, B.; Escher, P.; Tafuri, S.; Wang, M. *FEBS Lett.* **2000**, *484*, 159.
18. Cui, J.; Heard, T. S.; Yu, J.; Lo, J. L.; Huang, L.; Li, Y.; Schaeffer, J. M.; Wright, S. D. *J. Biol. Chem.* **2002**, *277*, 25963.
19. Laskowski, R. A. *J. Mol. Graph.* **1995**, *13*, 323.