RESEARCH ARTICLE

 View Article Online
 View Journal | View Issue

 Cite this: *Med. Chem. Commun.*,
 2019, 10, 1412

Dissecting the allosteric FXR modulation: a chemical biology approach using guggulsterone as a chemical tool†

 Daniela Passeri,^{‡a} Andrea Carotti,^{‡b} Jose M. Ramos Pittol,^{‡c} Gianmario Ciacchioli,^a
 Roberto Pellicciari,^a Saskia W. C. van Mil^{‡cd} and Antimo Gioiello^{‡*b}

Guggulsterone is a promiscuous ligand for endocrine and metabolic lipid receptors traditionally used to treat a number of diseases including diabetes, hyperlipidemia, atherosclerosis, and osteoarthritis. Although relatively weak, its activity at the farnesoid X receptor (FXR) is particularly intriguing as guggulsterone acts as an antagonist with a peculiar ability of gene selective modulation. We report here a chemical biology study with the aim to further characterize the biological action of guggulsterone at the FXR and to obtain further insights into the functional role played by noncanonical FXR binding pockets S2 and S3. Our results suggest that the FXR accessory pockets might act as potential targets for small molecules able to modulate the metabolic activation of the receptor without affecting the anti-inflammatory activity thus revealing a new approach for disclosing selective FXR modulators that might bypass potential side-effects from chronic treatments.

 Received 8th May 2019,
 Accepted 16th June 2019

DOI: 10.1039/c9md00264b

rsc.li/medchemcomm

1. Introduction

4,17(20)-Pregnadiene-3,16-dione, also known as guggulsterone (GS) (1) (Fig. 1), is the main active steroidal component of guggulipid. Guggulipid is a herbal extract of the gum resin of *Commiphora mukul* used in Ayurvedic medicine for more than 2000 years to treat inflammatory conditions and lipid-related disorders.^{1,2} Most of the pharmacological effects of guggulipid have been ascribed to the ability of GS to regulate the action of a subset of endocrine and metabolic nuclear receptors,³ including the progesterone receptor (PR), androgen receptor (AR), estrogen receptor (ER), glucocorticoid receptor (GR), mineralocorticoid receptor (MR), pregnane X receptor (PXR) and farnesoid X receptor (FXR). In particular, GS was reported to be an antagonist of the AR, MR and GR in the micromolar range, and to act as a partial PR agonist and a weak modulator of the FXR.

Among the activities of GS at nuclear receptors, the regulation of FXR activity is certainly an intriguing one.^{4–7} GS does not activate the FXR in a transactivation assay, while it

strongly inhibits the activation of the receptor promoted by the endogenous ligand chenodeoxycholic acid (CDCA) (2) (Fig. 1) in concentrations above 10 μ M. In cell-free assays, GS is unable to recruit a coactivator peptide, whereas it reverts the recruitment of the coactivator peptide by CDCA. Therefore, GS was originally classified as an FXR antagonist.^{4,5} However, additional experimental evidence indicated GS as a selective bile acid receptor modulator (SBARM) being able to selectively interfere with FXR target gene expression. Indeed, despite being characterized as a pure FXR antagonist in the coactivator association assay and shown to decrease the expression of bile acid activated genes, 1 was shown to enhance the FXR-agonist-induced expression of the bile salt export pump (BSEP), one of the FXR target genes.⁶

Fig. 1 Structures of guggulsterone (1) and bile acid-based agonists 2–4.

^a TES Pharma, Corso Vannucci 47, Perugia, Italy

^b Department of Pharmaceutical Sciences, University of Perugia, Perugia, Italy.
 E-mail: antimo.gioiello@unipg.it

^c Center for Molecular Medicine, UMC Utrecht, Utrecht University, Utrecht, The Netherlands

^d Tytgat Institute for Liver and Intestinal Research, Amsterdam UMC, Amsterdam, The Netherlands

[†] Electronic supplementary information (ESI) available. See DOI: 10.1039/c9md00264b

[‡] These authors equally contributed to the manuscript.

Molecular modelling studies aimed at defining the binding mode of **1** on the FXR provided first insights that could explain the peculiar action of GS (**1**) on the receptor. Next to docking into the canonical binding site (S1), **1** also docked into a novel binding pocket ('back door', S2), localized near the loop region between helix 1 and helix 2.⁸ This S2 pocket was found to be an alternative and energy favourable binding position.⁹ Notably, the existence of the S2 accessory site was recently validated by Yang and collaborators by a hydrogen/deuterium exchange mass spectrometry study.¹⁰

In line with these observations, docking experiments and molecular dynamics simulations performed on bile acid derivatives characterized by an extended carbamate-based side chain supported the accessibility of the S2 accessory pocket along with an additional binding area surrounding the canonical S1 binding site (the S3 pocket).¹⁰ While the S2 'back door' site is composed of Y270, F311 and T396 (NCBI Reference Sequence: NP_001193921.1), S3 is mostly exposed to the solvent, with only three residues (N281, K348 and S355) being the potential anchoring points. We hypothesize that the occupancy of these pockets induces conformational changes in the receptor that might be responsible for partial agonism and gene selective modulation. Therefore, understanding the molecular basis for the binding mode of **1** at these allosteric pockets may open the door to the rational design of unprecedented, gene selective FXR modulators.¹²

Herein, we report our ongoing efforts aimed at characterizing the peculiar biological activity of GS (**1**) with a focus on the FXR. In particular, we focused our attention towards the understanding of the functional role played by accessory binding pockets S2 and S3 in the FXR activity by biochemical and cell-based assays using GS (**1**) as a chemical tool in the presence of potent bile acid-based agonists such as OCA (**3**) and its carbamate derivative UPF-930 (**4**) (Fig. 1). We also investigated the transcriptome-wide FXR specific effects of GS (**1**) using wild type and FXR knockout mouse liver organoids as models. Furthermore, in order to elucidate the complex mechanisms of the FXR modulation through S2/S3 interaction, mutagenesis was performed based on the docking analyses and molecular dynamics simulations mentioned above.

2. Results

2.1 Nuclear receptor activity

Initially, the capacity of GS (**1**) to modulate a panel of nuclear receptors involved in metabolic and inflammatory pathways was assessed by AlphaScreen technology. We observed that **1** was able to activate the LXR β , PXR and PR at micromolar concentrations (Table 1), whereas it failed to induce activation of other nuclear receptors including the FXR. In an antagonist assay, GS (**1**) resulted in a broader action with IC₅₀ values ranging from 0.32 to 62 μ M (Table 1). Specifically, in line with previous findings,⁶ GS (**1**) was able to reduce the CDCA-dependent FXR activation with an IC₅₀ value of 25 μ M.

Table 1 Agonist and antagonist activity of guggulsterone (**1**) on nuclear receptors^a

Nuclear receptor	FXR	LXR α	LXR β	MR	PPAR α
EC ₅₀ (efficacy) ^b	>100	>100	3.6 (40%)	>100	>100
IC ₅₀	25	>100	>100	0.32	20
Nuclear receptor	PPAR δ	PPAR γ	PXR	RAR	RXR α
EC ₅₀ (efficacy) ^b	>100	>100	2.5 (155%)	>100	>100
IC ₅₀	>100	27	>100	>100	33
Nuclear receptor	VDR	ER	GR	PR	TR
EC ₅₀ (efficacy) ^b	>100	>100	>100	1.9 (115%)	>100
IC ₅₀	2	62	>100	>100	>100

^a EC₅₀ and IC₅₀ are expressed in μ M. Reference compounds and relative potency: FXR: CDCA EC₅₀: 8.0 μ M; LXR α : T091317 EC₅₀: 0.12 μ M; LXR β : T091317 EC₅₀: 0.06 μ M; MR: aldosterone EC₅₀: 0.0013 μ M; PPAR α : GW7647 EC₅₀: 0.003 μ M; PPAR δ : GW0742 EC₅₀: 0.004 μ M; PPAR γ : GW1929 EC₅₀: 0.012 μ M; PXR: SR12183 EC₅₀: 0.06 μ M; RAR: ATRA EC₅₀: 0.001 μ M; RXR α : 9-*cis*-retinoic acid EC₅₀: 0.004 μ M; VDR: dihydroxy-vitamin D EC₅₀: 0.0001 μ M; ER: estradiol EC₅₀: 0.001 μ M; GR: budesonide EC₅₀: 0.002 μ M; PR: progesterone EC₅₀: 0.05 μ M; TR: T3 EC₅₀: 0.0001 μ M. ^b Efficacy was calculated at the $E_{\max}$ of **1** curves versus the $E_{\max}$ of the positive control.

2.2 RNAseq data in organoids

In order to gain a better understanding of the FXR-dependent actions of GS (**1**) in a controlled physiologically relevant system, wild type and FXR knockout liver organoids were treated with 50 μ M GS or a vehicle (DMSO). An RNAseq analysis revealed that upon stimulation with GS (**1**), a large number of genes are up and down-regulated both in the organoids derived from wild type or FXR^{-/-} mice, confirming its ability to regulate not only the FXR but also other pathways. Overall, the treatment with GS (**1**) led to both activation and repression of genes under regular organoid culturing conditions. However, GS (**1**) mediated FXR-driven effects are largely repressive or inhibitory (black clusters in Fig. 2A), in line with the observations in different systems.^{4,13}

In order to contextualize the nature and effect of the FXR mediated effects, we performed an upstream regulator analysis (IPA) on the FXR gene subgroup (Fig. 2C). From the significantly enriched pathways, we found that inflammation related factors have the highest changes in the Z-score. In this context, the Z-score correlates to the activation or repression of the analysed pathway. Altogether, these results highlight that the FXR mediates a notable amount of GS (**1**) effects in the modulation of the NF κ B pathway. This was also reflected on the genes with the highest down-regulation in the wild type vs. the FXR^{-/-} (Fig. 2D).

2.3 Cell-based and biochemical competitive assays

We next evaluated the response upon co-administration of GS (**1**) with efficacious FXR agonists **3** and **4** in both biochemical and cell-based assays. To assess the antagonistic activity of **1**, a reporter assay was performed. HEK293T cells

Fig. 2 FXR-dependent effects of guggulsterone (1). (A) Heatmap of RNA-sequencing data of two independent wild type and FXR^{-/-} liver organoid preparations upon 50 μ M guggulsterone (1) or DMSO stimulation for 8 hours. Results are shown as log₂-fold induction of guggulsterone compared to DMSO. Ward.D clustering with tree cut ($k = 6$) is depicted in black/white bars. Differential clusters between genotypes are colored in black. (B) Venn diagram of significantly regulated genes in wild type or FXR^{-/-} organoids, FDR < 0.05. (C) Ingenuity pathway analysis for upstream regulators of FXR-driven guggulsterone mediated target genes. (D) log₂-fold induction of top FXR-dependent upregulated and downregulated guggulsterone target genes.

were transfected with the FXR, RXR and a reporter plasmid and incubated with OCA (3) or the carbamate derivative UPF-930 (4) in the presence or absence of 1. As expected, GS (1) was able to antagonize the FXR ligand-mediated activation (Fig. 3).

To further explore the mechanism by which GS (1) antagonizes the bile acid-mediated FXR activation, a biochemical competitive assay was performed by means of AlphaScreen technology. Remarkably, when tested in combination with 3 and 4, GS (1) acted as a non-competitive antagonist (Fig. 4). This effect was likely due to the ability of GS (1) to bind at the S2 pocket when the canonical binding site is occupied by agonists 3 or 4. Compound 4 in particular is characterized by a steroidal scaffold that occupies the canonical binding pocket S1 and by a flexible side chain that can bind both the S2 and S3 accessory pockets (see Fig. 6).¹¹

Fig. 3 Transactivation assay performed using the human FXR combined with the reporter vector harbouring the canonical human FXR IR1 response element sequence repeated three times. Guggulsterone (1) is able to antagonize the FXR activation in combination with OCA (3) or UPF-930 (4) at 0.5 μ M (left and right panel, respectively; $P < 0.05$). Data represent mean values $\pm$ SD of at least three independent experiments. * $p < 0.05$ vs. OCA or UPF-930 alone activity.

2.4 Mutagenesis study

Docking analysis and molecular dynamics simulations were performed on apo- and holo-forms of the FXR to shed light on the complex mechanism of receptor modulation by S2/S3 ligand interaction. Starting from the CDCA FXR ligand binding domain complex X-ray structure (Protein Data Bank ID 4QE6), a docking run was carried out on steroids 1, 3 and 4 according to the computational workflow depicted in Fig. 5.

Ligand docking analysis resulted in the identification of three key residues (FXR α 4 isoform residue numbering) predicted to be involved in the ligand binding interaction of 1 and the carbamate 4: two located in the S2 binding pocket (T396 and Y270) and one located in the S3 binding pocket (S355) (Fig. 6). Next, five virtual models of FXR mutants (Y270E, S355E, S355F, T396E and T396F) were generated and evaluated for their ability to affect the binding of the selected ligands, when compared with the FXR wild type. Only compound 4 displayed an S3 putative binding pose, therefore it

Fig. 4 AlphaScreen assay performed using the human FXR ligand binding domain and the human coactivator SRC1. Guggulsterone (1) shows a non-competitive binding profile when tested in combination with OCA (3) or UPF-930 (4). Data represent mean values $\pm$ SD of at least three independent experiments.

Fig. 5 Computational protocol applied to identify, virtually mutate and study the interacting residues in the S2 and S3 sites. The workflow resulted in the selection of T396E, S355E and S355F mutations to be expressed and profiled through biochemical assays.

was the only compound dynamically simulated with the S355 set of mutants.

From the molecular dynamics analyses, the T396E, S355E and S355F mutants were selected for validation in the cellular

Fig. 6 FXR (human $\alpha 4$ isoform residue numbering) residues virtually mutated in the accessory binding pockets S2 (Y270 and T396, yellow) and S3 (S355, cyan), predicted to be involved in the binding of **1** (yellow) and **4** (cyan and blue, in S2 and S3 putative binding modes, respectively). **3** is shown in green. Helix 12 is highlighted in magenta. S1, S2 and S3 binding sites are indicated by green, yellow and cyan transparent forms, respectively.

assays. Indeed, with respect to the wild type model, the T396E mutant dynamically highlighted a change in the interaction energy of GS (**1**) with the FXR ligand binding domain towards more favourable values (see the ESI† for details). The S3 mutants were chosen to gain insights into the *in silico* observed possibility of **4** to alternatively bind the S2 and S3 sites (Fig. 6). Thus, HEK293T cells were transfected with different expression vectors harbouring the FXR wild type or mutants. Remarkably, the presence of the mutation did not affect the expression of the FXR proteins (Fig. S1, ESI†).

Next, the cells were treated with OCA (**3**) or UPF-930 (**4**) alone and in combination with GS (**1**) to study the interaction between GS and the selected amino acid on the S2 and S3 binding pockets. We found that GS (**1**) significantly inhibited the activation of the FXR by full agonists **3** and **4** on both the wild type and T396E mutant, with an enhanced antagonist effect observed for an S2 mutated receptor form (Fig. 7). Furthermore, T396E mutation reduced the efficacy of both agonists thus indicating that both the occupation and conformation of the S2 accessory pocket influence the agonists binding at the canonical binding site S1.

The transactivation ability showed no difference for the S355E and S355F mutants relative to the wild type receptor. Unfortunately, these results did not clearly confirm the binding of **4** to the S3 site, but rather raised another hypothesis to be investigated. Namely, the S355 residue is located on the loop between helix 5 and helix 6, and the S3 binding site as a whole is relatively exposed and open, thus favouring the rearrangement of both the protein residues and the ligand conformational disposition of the carbamate side chain. Overall, these results evidenced that **1** was likely to antagonize the FXR transactivation ability mediated by both **3** and **4** *via* the S2 pocket and that the amino acid T396 was the key residue involved in the binding with **1**.

Fig. 7 Transactivation assay on HEK293T cells performed by transient transfection of the full-length FXR wild type or mutated FXR with the canonical FXRE containing 3 inverted repeats (IR1). The FXR S2 mutant T396E shows a decreased transactivation ability upon co-administration of OCA (**3**) or UPF-930 (**4**) and GS (**1**), whereas no differences were observed on the S355 mutants. Data represent mean values $\pm$ SD of at least three independent experiments. * $p < 0.05$ versus FXR wild type activity.

To assess whether the mutation at the S2 pocket may affect the potency and/or efficacy of **1**, the cells were first transfected with the FXR wild type or T396E mutant and then stimulated with agonist **3** or **4** alone, and in a full dose-response of GS (**1**) *versus* both agonists (**3** and **4**) (Fig. 8). The results showed that the T396E mutation had a minor impact on the affinity (EC_{50}) of the tested compounds (Fig. 8A and B), but seemed to have a major influence on their efficacy suggesting a possible allosteric effect of the S2 pocket on the canonical pocket S1. In particular, the reduction of the efficacy appeared more evident for carbamate **4** whose extended side chain was shown to occupy the S2 pocket (Fig. 6). The dose-response experiment of **1** in the presence of both agonists demonstrated that the T396E mutation outperformed the antagonism effect on the FXR wild type, in contrast no significant changes were observed for the relative potency (Fig. 8C and D).

3. Discussion and conclusions

Allosteric modulation of FXR is still far from being a validated approach to target and modulate the activity of the receptor. Evidence of such mechanisms in nuclear receptors is, however, reported in the literature¹⁴ and has offered novel directions for designing new chemical probes and efficacious lead compounds on the route of novel drugs.

Fig. 8 Transactivation assay on HEK293T cells transfected with the FXR wild type or mutated FXR (T396E). (A and B) Dose-response experiments of **3** and **4**. The FXR S2 mutant T396E decreased the transactivation efficacy of **3** and **4** with no major changes on the EC_{50} values. (C and D) Dose-response experiments of **1** in the presence of **3** at a fixed concentration of 0.5 μ M and **4** at 1 μ M. The FXR S2 mutant T396E increases the efficacy of **1** in the presence of both agonists. The efficacy of **1** was calculated at 10 μ M. Data represent mean values $\pm$ SD of at least three independent experiments.

In this article, we aim to provide our contribution to this open question and define a new working hypothesis for future research programs in the field of the FXR. Starting from our previous findings^{8,9,11} and recent reports,¹⁰ we decided to better characterize the activity of GS (**1**) at the FXR and to employ this naturally-occurring steroid as a chemical tool for exploring the relevance of the receptor accessory pockets S2 and S3.

Originally identified as an FXR antagonist *in vitro*,^{4,5} GS (**1**) was later classified as a SBARM due to its ability to selectively interfere with FXR target genes.⁶ Experimental evidence supports a model in which *in vitro* antagonism and gene selective modulation of the FXR by GS (**1**) might be due to the binding of the molecule at the receptor allosteric pockets with a more favourable binding energy than the canonical S1 pocket.^{8,9} This finding has provided new clues to explain the peculiar action of GS at the FXR and has delineated the intriguing possibility of modulating the FXR through these allosteric sites.

Herein, we have performed computational modelling and biochemical and biological assays combined with mutagenesis experiments so as to elucidate the biological activity and molecular mechanisms of the action of **1** at the FXR. In particular, the scope of the study was to provide evidence that the binding of **1** at the noncanonical pockets of the receptor may mediate the antagonism toward agonist binding at the canonical S1 binding site and/or to the S1 and S2 allosteric sites as for the extended side chain bile acid derivative UPF930 (**4**). Firstly, the activity of GS (**1**) on a panel of 15 nuclear receptors was assessed both in the agonist and antagonist modes (Table 1). While this represents a comprehensive profiling of the compound on major human transcription factors, the data confirmed the promiscuous nature of **1** as a ligand. In particular, GS (**1**) was unable to activate the FXR up to 100 μ M and to inhibit the activity of the endogenous ligand CDCA (**2**) with an IC_{50} value of 25 μ M. The transactivation assay showed that **1** was also able to inhibit the FXR ligand mediated activation upon binding of the potent agonists, bile acid derivatives **3** and **4** (Fig. 3). We have speculated that this effect can be ascribed to the binding of **1** to the S2 accessory pocket in view of the non-competitive profile of GS (**1**) when tested in combination with **3** and **4** (Fig. 4).

To confirm the hypothesis and to pinpoint the S2 amino acids that may affect the binding of **1**, a computational study was performed leading to the identification of the T396 residue as a potential critical hotspot. In particular, the T396E molecular dynamics in complex with GS (**1**) highlighted a higher occupancy of the S2 accessory pocket with respect to the wild type trajectory. Indeed, **1** moved further inside the S2 area by engaging an H-bond with R341 located at the entrance of the S1 canonical binding pocket. Notably, R341 is a key residue involved in a stable interaction with the C24 tail moiety (*i.e.* hydrogen acceptor oxygen) of bile acid derivatives, particularly strong (charge-charge type), in the case of the C24-carboxylic group as in CDCA and OCA (pdb codes 1OSV, 6HL1).¹⁶ This working hypothesis was also investigated by a

transactivation mutagenesis study. The results indicated that **1** exerted a more potent antagonism effect on the S2 T396E mutant compared to the wild type (Fig. 7). Importantly, a full dose-response experiment of GS (**1**) *versus* both agonists **3** and **4** on the T396E mutant was likely to affect the efficacy rather than the potency of the tested compounds (Fig. 8). This effect was particularly relevant for carbamate **4** whose extended side chain is supposed to interact with the S2 allosteric pocket.^{8,11}

In search for a better definition of the binding mode of **4**, especially at the S3 accessory pocket, S355E and S355F mutants were chosen according to the molecular dynamics analyses. Unfortunately, the S355E and S355F mutant tests led to uninformative results, showing no differences when compared to the wild type form. Similarly, no differences were observed among the FXR wild type and S2/S3 mutants in the transrepression of the TNF α /NF κ B mediated inflammatory response (data not shown).

Next, we performed an RNAseq analysis of organoids treated with **1** in the absence of agonists to address the potential of the allosteric binding. In our experimental setting, the effects of **1** on FXR dependent targets were mostly inhibitory or repressive. Despite the absence of agonists, this result provides strong evidence for the anti-inflammatory effect of **1** by the FXR and the promiscuity of the compound as a nuclear receptor multitarget binder (Table 1).

In summary, our data suggest that the FXR accessory pockets might act as potential targets for small molecules to selectively interfere with the metabolic action of the FXR. In particular, our findings reveal a novel strategy for the design and development of a new class of selective FXR modulators that might bypass the potential side-effects from chronic treatments. Virtual and fragment-based screening campaigns, as well as the rational design of steroid molecules, present possible entries to specifically target the FXR allosteric sites and identify the allosteric modulators. Not less important, our study provides clues to reconsider the molecular basis of activation and related biological properties of FXR modulators that have shown to be endowed with a partial agonist and gene-selective modulation effect.

4. Experimental section

4.1 Chemistry

Guggulsterone (**1**) was prepared according to a previously reported synthesis.¹⁵ The correct stereochemistry (*E*-isomer) and purity (>98%) were obtained by NMR spectroscopy and HPLC analysis.

4.2 Computational chemistry

All the calculations were performed with the Schrödinger Suite 2019-1 (Schrödinger Inc., NY, USA). The crystal structure of the CDCA FXR ligand binding domain (Protein Data Bank ID 4QE6) was downloaded and prepared with the default settings of the Protein Preparation Wizard tool of Maestro 11.9 (Schrödinger Inc.). The grid has been generated

with Glide 8.2, with the centre located on the centre of mass of the crystallized ligand. The inner grid box was sized 20 × 20 × 20 Å to also include the S2 and S3 sites. Ligands have been prepared with the LigPrep module generating all ionization states at pH 7 ± 2. Docking studies have been carried out using a Glide 8.2 standard precision (SP) method and the five best poses of each molecule were stored as output. The poses have been visually inspected and selected for the generation of the molecular dynamics simulation systems. In particular, for GS (**1**), the best ranked pose interacting in S2 was chosen. For OCA (**3**), the best pose reproducing the co-crystallized pose (Protein Data Bank ID 1OSV, rat FXR)¹⁶ was selected. For UPF-930 (**4**), two poses were picked, one orienting the carbamate side chain in the S2 site and another in the S3. All these poses, together with the apo forms of the wild type and mutated (Y270E, S355E, S355F, T396E and T396F in α 4 isoform residue numbering) FXR ligand binding domain, have been used as starting structures to build the molecular dynamics systems (following the Fig. 4 scheme). All the 20 models were subjected to a molecular dynamics simulation (MD) analysis carried out with the Desmond package. In particular, the complexes were firstly solvated with water TIP3P and neutralized with the appropriate counterions using the system builder module. The box used was orthorhombic, with each side having a minimum distance of 10 Å from any atom of the complex. The force field used was OPLS3e and all the other options were set to default values. The dynamic protocol was composed of subsequent steps, with the first being a relaxing minimization, then a progressive increase of temperature to reach 300 K in 140 ps, and finally a production run of 100 ns. Frames of the simulation were recorded every 50 ps, resulting in a trajectory of 2000 frames. The MD trajectories were then analysed using the Desmond tools, taking into account two parameters: the backbone RMSD and the interaction energy between the ligand and protein (for the **1** and **4** holo systems). The former was functional to check the protein stability, and the latter was important to evaluate the effect of the mutations in the S2/S3 sites. The mutations suggesting a high impact on the S2/S3 binding sites were T396E, S355E and S355F.

4.3 Biology

4.3.1 AlphaScreen assay. Activity on the FXR and other nuclear receptors was assayed by using a recruitment coactivator assay, namely AlphaScreen technology, according to the manufacturer's instructions. In brief, the assay was performed in a white, low-volume, 384-well OptiPlate (PerkinElmer Life and Analytical Sciences, Waltham, MA) using a final volume of 25 μ L containing 10 nM glutathione transferase-tagged FXR-LBD protein and 30 nM biotinylated Src-1 peptide. The stimulation was carried out with 1 μ L of ligand (solubilized in 100% DMSO) for 30 min at room temperature, and the luminescence was read using an EnVision microplate analyzer (PerkinElmer Life and Analytical Sciences) after incubation with a detection mix (acceptor and

donor beads) for 4 h at room temperature in the dark. A specific positive control was used for each nuclear receptor tested as described in the results.

4.3.2 Cell cultures. HEK 293T/17 cells were cultured in Dulbecco's modified Eagle's medium high-glucose (DMEM) supplemented with 10% FBS, 100 units per ml penicillin and 100 $\mu\text{g ml}^{-1}$ streptomycin sulfate. The cells were grown at 37 °C in an atmosphere plus 5% CO_2 .

4.3.3 Immunoblot analysis. Total protein extracts were obtained using RIPA lysis buffer and protein concentration was assayed according to the Bradford method. The protein samples were denatured by heating at 90 °C in 1× final Laemmli buffer and resolved by electrophoresis on 4–15% Mini-PROTEAN TGX gels. After the protein transfer to the nitrocellulose, membranes were probed with rabbit polyclonal Ab anti-FXR (FXR H-130) sc-13063, chicken polyclonal Ab anti-histone H4 (Ab134213), goat anti-rabbit IgG (H + L)-HRP conjugate Bio-Rad 1706515 and goat anti-chicken IgY H&L (HRP) (ab97135). The proteins were then visualized by chemiluminescence using a Pierce ECL western blotting substrate with Bio-Rad ChemiDoc imaging systems and imaging quantification was carried out with ImageJ.

4.3.4 Transient transfection and transactivation. For the transactivation assay, HEK293T cells were obtained from the American Type Culture Collection (ATCC-CRL-11268). The cells were transiently transfected with 100 ng of reporter vector pGL4.23 FXR-IR1x3, 10 ng of expression plasmid pCDNA3.1(+) hFXR CDS and 10 ng of pSG5 RXR α . 40 ng of pGL4.74 was used as an internal control for the transfection efficiency. The pGEM3 vector was added to normalize the amounts of DNA transfected in each assay (2 μg). All transfections were performed using a FuGENE HD ratio of 5:2 (Roche, Mannheim, Germany) according to the manufacturer's protocol. Twenty-four hours after transfection, the cells were stimulated with the indicated compounds for a further 18 h. Control cultures received a vehicle (0.1% DMSO) alone. The luciferase values were normalized with *Renilla reniformis* luciferase units, for the transfection efficiency.

4.3.5 Mutagenesis. Mutagenesis on pCDNA3.1(+) hFXR CDS to generate its T396E mutant was performed using an Agilent QuikChange II Site-directed mutagenesis kit and specific primer design was conducted using an Agilent QuikChange Primer Design tool <https://www.agilent.com/store/primerDesignProgram.jsp>. T396E forward primer 5'-gactcaagaggagtagtctgtctgaggcaattgttatctgtctc-3'; T396E reverse primer 5'-gagacagataacaattgctctcaagcagagcactactctcttgagtc-3'.

4.3.6 Organoid culture. Liver organoids from the wild type and FXR $^{-/-}$ mice were generated, expanded and differentiated as previously described.¹⁷ These organoids were differentiated into the hepatocyte lineage before the drug treatment and treated with either DMSO or 50 μM GS for 8 hours before the RNA extraction.

4.3.7 mRNA isolation and sequencing. The RNA was isolated using a Trizol® reagent (Ambion/Life Technologies). Libraries were generated using a NEXTflex™ Rapid RNA-seq kit (Bio Scientific) using 1 μg total RNA, and sequenced by a

Nextseq500 platform (Illumina, Single-end, 75 bp reads). The obtained sequences were aligned to the mouse genome mm9 using TopHat2 (ref. 18) and the differential expression of genes was calculated using Cufflinks.¹⁹ Heatmap generation and hierarchical clustering of the significantly changed genes were generated with the BioConductor package in Rstudio.²⁰ The values shown are the fold induction between GS and DMSO treatment.

4.3.8 Pathway analysis. The pathway analysis was performed using Ingenuity Canonical Pathway Analysis (IPA; Ingenuity® Systems, www.ingenuity.com) on the gene subsets defined by their FXR derived responsiveness, following the clustering analysis. The IPA library of canonical pathways was used as a reference and the top-ranked results were selected by considering the differences in Z-scores between the wild type and FXR $^{-/-}$.

Conflicts of interest

RP and AG are co-founders of TES Pharma.

Acknowledgements

This work could not have been undertaken without the financial support from TES Pharma (Perugia, Italy) and Intercept Pharmaceuticals (New York, USA). This work was also supported by FP7 Marie Curie Actions IAPP (FXR-IBD, 611979). NVIDIA Corporation is gratefully acknowledged for the donation of the GPU used to perform the molecular dynamics simulations. We also thank Nico Lansu and Noortje van den Dungen for the preparation of sequencing libraries at the Epigenomics facility, UMC Utrecht, and the Utrecht Sequencing Facility for providing sequencing service and data.

Notes and references

- 1 G. V. Satyavati, *Econ. Med. Plant Res.*, 1991, 5, 47.
- 2 N. L. Urizar and D. D. Moore, *Annu. Rev. Nutr.*, 2003, 23, 303.
- 3 T. P. Burris, C. Montrose, K. A. Houck, H. E. Osborne, W. P. Bocchinfuso, B. C. Yaden, C. C. Cheng, R. W. Zink, R. J. Barr, C. D. Hepler, V. Krishnan, H. A. Bullock, L. L. Burris, R. J. Galvin, K. Bramlett and K. R. Stayrook, *Mol. Pharmacol.*, 2005, 67, 948.
- 4 N. L. Urizar, A. B. Liverman, D. T. Dodds, F. V. Silva, P. Ordentlich, Y. Yan, F. J. Gonzalez, R. A. Heyman, D. J. Mangelsdorf and D. D. Moore, *Science*, 2002, 296, 1703.
- 5 J. Wu, C. Xia, J. Meier, S. Li, X. Hu and D. S. Lala, *Mol. Endocrinol.*, 2002, 16, 1590.
- 6 J. Cui, L. Huang, A. Zhao, J. L. Lew, S. Sahoo, P. T. Meinke, I. Royo, F. Pelaez and S. D. Wright, *J. Biol. Chem.*, 2003, 278, 10214.
- 7 E. Owsley and Y. L. Chiang, *Biochem. Biophys. Res. Commun.*, 2003, 304, 191.
- 8 R. Pellicciari, A. Gioiello, G. Costantino, B. M. Sadeghpour, G. Rizzo, U. Meyer, D. J. Parks, A. Entrena-Guadix and S. Fiorucci, *J. Med. Chem.*, 2006, 49, 4208.

- 9 U. Meyer, G. Costantino, A. Macchiarulo and R. Pellicciari, *J. Med. Chem.*, 2005, **48**, 6948.
- 10 L. Yang, D. Broderick, Y. Jiang, V. Hsu and C. S. Maier, *Biochim. Biophys. Acta*, 2014, **1844**, 1684.
- 11 A. Gioiello, A. Macchiarulo, A. Carotti, P. Filipponi, G. Costantino, G. Rizzo, L. Adorini and R. Pellicciari, *Bioorg. Med. Chem.*, 2011, **19**, 2650.
- 12 V. Massafra, R. Pellicciari, A. Gioiello and S. W. C. van Mil, *Pharmacol. Ther.*, 2018, **191**, 162.
- 13 J. Wu, C. Xia, J. Meier, S. Li, X. Hu and D. S. Lala, *Mol. Endocrinol.*, 2002, **16**, 1590.
- 14 F. A. Meijer, I. A. Leijten-van de Gevel, R. M. J. M. de Vries and L. Brunsveld, *Mol. Cell. Endocrinol.*, 2019, **485**, 20.
- 15 A. Gioiello, R. Sardella, E. Rosatelli, B. M. Sadeghpour, B. Natalini and R. Pellicciari, *Steroids*, 2012, **77**, 250.
- 16 L. Z. Mi, S. Devarakonda, J. M. Harp, Q. Han, R. Pellicciari, T. M. Willson, S. Khorasanizadeh and F. Rastinejad, *Mol. Cell*, 2003, **11**, 1093.
- 17 M. Huch, H. Gehart, R. van Bortel, K. Hamer and H. Clevers, *Cell*, 2015, **160**, 299.
- 18 D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley and S. L. Salzberg, *Genome Biol.*, 2013, **14**, R36.
- 19 C. Trapnell, D. G. Hendrickson, M. Sauvageau, L. Goff, J. L. Rinn and L. Pachter, *Nat. Biotechnol.*, 2013, **31**, 46.
- 20 W. Huber, V. J. Carey, R. Gentleman, S. Anders, M. Carlson, B. S. Carvalho, H. C. Bravo, S. Davis, L. Gatto, T. Girke, R. Gottardo, F. Hahne, K. D. Hansen, R. A. Irizarry, M. Lawrence, M. I. Love, J. MacDonald, V. Obenchain, A. K. Oleś, H. Pagès, A. Reyes, P. Shannon, G. K. Smyth, D. Tenenbaum, L. Waldron and M. Morgan, *Nat. Methods*, 2015, **12**, 115.