Impaired transitioning of the FXR ligand binding domain to an active state underlies a PFIC5 phenotype

Annika Behrendt[1], Jan Stindt[2], Eva-Doreen Pfister[3], Kathrin Grau[1], Stefanie Brands[1], Alex Bastianelli[4], Carola Dröge[2,4], Amelie Stalke[5], Michele Bonus[1], Malte Sgodda[6], Tobias Cantz[6,7], Sabine Franke[8], Ulrich Baumann[3], Verena Keitel[2,4]* and Holger Gohlke[1,9]*

[1] Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University, Düsseldorf, Germany

[2] Department of Gastroenterology, Hepatology and Infectious Diseases, Medical Faculty University Hospital Düsseldorf, Heinrich Heine University, Düsseldorf, Germany

[3] Pediatric Gastroenterology and Hepatology, Department for Pediatric Kidney, Liver and Metabolic Diseases, Hannover Medical School, Hannover, Germany

[4] Department of Gastroenterology, Hepatology and Infectious Diseases, Medical Faculty, University Hospital Magdeburg, Otto von Guericke University, Magdeburg, Germany

[5] Department of Human Genetics, Hannover Medical School, Hannover, Germany

[6] Research Group Translational Hepatology and Stem Cell Biology, Department of Gastroenterology, Hepatology, Infectious Diseases and Endocrinology, Hannover Medical School, Hannover, Germany

[7] REBIRTH-Research Center for Translational Regenerative Medicine, Hannover Medical School, Hannover, Germany

[8] Institute of Pathology, University Hospital Magdeburg, Otto von Guericke University, Magdeburg, Germany

[9] John-von-Neumann-Institute for Computing, Jülich Supercomputing Center, Institute of Biological Information Processing (IBI-7: Structural Biochemistry), and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, Jülich, Germany

*corresponding authors

Contact information / Correspondence

Prof. Dr. rer. nat. Holger Gohlke

Institute for Pharmaceutical and Medicinal Chemistry,

Heinrich-Heine-Universität Düsseldorf

Universitätsstr. 1

40225 Düsseldorf

Germany

Phone: +49 211 81 13662

Email: gohlke@uni-duesseldorf.de

Prof. Dr. med. Verena Keitel

Department of Gastroenterology, Hepatology and Infectious Diseases, Medical Faculty of Otto-von-Guericke University Magdeburg, University Hospital Magdeburg AöR

Leipziger Str. 44

39120 Magdeburg

Germany

Fax: +49 391 6713105

Phone: +49 391 6713100

Email: verena.keitel-anselmino@med.ovgu.de

Keywords: farnesoid X receptor, *NR1H4*, progressive familial intrahepatic cholestasis, missense variant, molecular dynamics simulations

Abbreviations

FXR farnesoid X receptor, LBD ligand binding domain; PFIC progressive familial intrahepatic cholestasis; BA bile acids; RXR retinoic X receptor; MD Molecular Dynamics; NR nuclear receptor; RMSD root mean square deviation; H12 helix 12; AF activation function; OCA obeticholic acid; CDCA chenodeoxycholic acid; NCoA2 nuclear receptor coactivator 2;

Funding

This study was supported by the BMBF through HiChol (01GM1904A and 01GM2204A to V.K. and C.D.; 01GM1904A and 01GM2204B to H.G.; 01GM1904B and 01GM2204C to U.B., A.S., E.P., TC).

Acknowledgment

Expert technical assistance by Paulina Philippski and Aileen Nötzold is gratefully acknowledged. We are grateful for computational support by the “Zentrum für Informations und Medientechnologie” at the Heinrich-Heine-Universität Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) to H.G. on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user IDs: VSK33; FIC1).

Conflict of interest statement

The authors declare that there is no potential conflict of interest.

Abstract

Nuclear receptor farnesoid X receptor (FXR) acts as a key regulator of bile acid pool homeostasis and metabolism. Within the enterohepatic circulation, reabsorbed bile acids act as agonists on FXR, which transcriptionally controls the synthesis and transport of bile acids. Binding occurs in the ligand binding domain (LBD), favoring a conformational change to the active state in which helix 12 interacts with the LBD to form an interaction surface for nuclear co-activators. The homozygous missense variant T296I, identified in a PFIC5 patient, is located close to the critical helix 12 interaction. Here, we identified reduced transcriptional activity of the variant protein on the downstream targets BSEP and SHP *in vitro* and within the patient's liver. Analysis of the structural dynamics of the conformational change from an inactive to an active state of the FXR LBD with molecular dynamics simulations revealed that while the wildtype protein frequently transitions into the active state, this movement and the necessary perfect placement of helix 12 was significantly impeded in the T296I mutated protein. To our knowledge, this is the first study to describe the conformational change from an inactive to an active state of the FXR LBD. This might be useful for new therapeutic approaches targeting the activation of FXR. Overall, combining *in vivo* data with *in vitro* and *in silico* experiments, we suggest a molecular mechanism underlying the PFIC phenotype of a patient with an FXR missense variant.

Introduction

Progressive familial intrahepatic cholestasis (PFIC) is a rare group of genetic disorders that affect the liver's ability to excrete bile constituents, resulting in impaired bile flow, subsequent intrahepatic cholestasis, and progressive liver damage and failure (1, 2).

Farnesoid X receptor (FXR), encoded by the *NR1H4* gene, is a nuclear receptor (NR) responsive to bile acids (BA) and a key regulator of BA metabolism, playing a pivotal role in maintaining BA homeostasis by controlling BA synthesis, transport, and detoxification (3, 4). *NR1H4* variants associated with PFIC (subtype 5) were characterized by coagulopathy and a rapid progression toward end-stage liver disease (5-7). While most patients carried bi-allelic protein-truncating variants (5-8), only two *NR1H4*-associated PFIC patients carrying homozygous missense variants have been identified (7, 9). While one patient died on the transplant waiting list due to end-stage liver disease at the age of 9 months (c.557G>A) (7), the other patient was successfully transplanted at the age of 8 months (c.887C>T, p.(Thr296Ile), referred to as T296I in the following) and is currently 10 years old (9). FXR and BSEP staining was found negative in the liver tissue of PFIC patients with protein-truncating NR1H4 variants (5, 8). To determine the contribution of the homozygous *NR1H4* T296I missense variant to the PFIC phenotype of our patient, we studied the localization and transcriptional activity of the mutated protein *in vitro*.

Molecular dynamics (MD) simulations have proven useful in elucidating the functional mechanisms of protein activity (10). In particular, nuclear receptors (NRs) have benefited from this in-depth analysis as their functions are often diverse, and subtle changes in ligands can lead to altered conformations and, thus, protein activity (11-13). The positioning of helix 12 (H12), forming part of the activation function 2 (AF2) surface, is pivotal for NR activity via the recruitment of coregulatory proteins. Coactivators interact with the AF2 surface using a conserved LXXLL motif (14), while antagonist-bound NRs favor corepressor binding to the AF2 surface with a larger hydrophobic motif and blocking the active positioning of H12 (15). Several MD studies of the LBD of FXR have underlined the importance of H12 positioning (16-18). However, the transitioning from the inactive to the active conformation as well as the effect of

single-site missense variants on the function of FXR has so far not been analyzed by MD studies. Thus, we employed MD simulations to analyze the conformational change from an inactive to an active state and evaluated the impact of the T296I variant both with a localized distance measurement and with regard to its influence on H12 positioning.

Materials and Methods

Plasmids, cloning and mutagenesis

The BSEP promoter plasmid based on pGL3-basic ($BSEP^{prom}$-Luc) was a kind gift from Roche. The human *SHP* promoter (bases -572 to +10, GenBank Accession Number AF044316) (19) was amplified by PCR from a healthy human liver genomic DNA pool. DNA sequencing was performed for all cDNAs used (Eurofins). Note that the numbering of the protein variant (T296I) is based on the alpha1 isoform (Uniprot acc. Q96RI1-1). For details on the cloning strategies see SI.

Immunofluorescence staining of HEK293 cells

HEK293 cells seeded onto glass coverslips in 12-well plates were transiently transfected with 1µg each of FXR and retinoic X receptor (RXR) α expression constructs for 48h. After 24h, cells were stimulated with obeticholic acid (OCA, INT-747, 10µM) and 9-cis-retinoic acid (9-cis-RA, 1µM). Cells on coverslips were washed with PBS before fixation with ice-cold methanol (30sec). After blocking in UltraVision protein block (ThermoFisher Scientific) for 30min, cells were stained for 1h at 1:100 with rabbit anti-FXR (H-130; sc-13063, Santa Cruz Biotechnology) followed by staining at 1:250 with goat anti-rabbit-IgG-FITC (Jackson ImmunoResearch) and DAPI at 1:20.000. Coverslips were mounted on microscopic slides using Dako fluorescence mounting medium.

Western Blot

HEK293 cells seeded into 6-well plates were transiently transfected for 48h with 2µg per well of either WT or mutant *pno*Cherry-FXR as described above and in SI. Membranes were blocked with 5% BSA in TBS-T for 1h before overnight incubation with rabbit anti-FXR (H-130; 1:2.000) and mouse anti-β-actin (ab6276, Abcam; 1:10.000) followed by incubation with goat anti-rabbit-IgG-AlexaFluor 647 and goat anti-mouse-IgG-AlexaFluor 488 (both at 1:5.000). Fluorescent signals were detected using a ChemiDoc MP imaging system (Biorad).

Luciferase Assay

Luciferase assays were performed using the Dual Luciferase reporter assay (Promega) according to the manufacturer´s instructions. Briefly, HEK293 cells kept in DMEM containing 10% fetal calf serum (FCS) were seeded onto 12-well plates at 150.000 cells per well and transfected the next morning with 1µg of the $BSEP^{prom}$-Luc or SHP^{prom}-Luc plasmid and 100ng each of FXR and RXR expression plasmids using Fugene HD (Promega) at a ratio of 2.5:1 (reagent:DNA). Cells transfected with FXRα1/2 expression plasmid were stimulated with 10µM OCA (INT-747), cells transfected with RXRα were stimulated with 1µM 9-cis-retinoic acid (9-cis-RA), cells transfected with both FXRα1/2 and RXRα expression plasmids were stimulated with both ligands.

RNA preparation, reverse transcription, pre-amplification, and PCR analysis

Total RNA was extracted and purified using the AmoyDx FFPE DNA/RNA Kit, (Amoy Diagnostics Co.) according to the manufacturer's instructions. 100ng of RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). After pre-amplification, qPCR was carried out with different TaqMan™ Gene Expression Assays (Applied Biosystems). Relative quantification of mRNA was performed according to the comparative $2^{-\Delta\Delta CT}$ method with SDHA as an endogenous control (see SI for detailed information).

Structure modeling and molecular dynamics simulations

To analyze the impact of the variant T296I, the ligand binding domain (LBD) structure of FXR (Q96RI1-1, residues 248 to 476) was modeled based on the chenodeoxycholic acid (CDCA)- and NCoA2 peptide-bound X-ray crystal structure of the FXR LBD (PDB ID 6HL1) (13), representing the active state of FXR, using SWISS-MODEL (20). To model the inactive state with H12 not interacting with the LBD core, the loop between helix11 and H12 was remodeled within PyMOL (Schrödinger, LLC, New York). In detail, residues 460 to 466 (^{460}VNDHKFT^{466}) were removed and readded, pointing away from the LBD core, followed by the α-helix H12.

MD simulations were performed for both the active and inactive states in the presence of the endogenous ligand CDCA and a short peptide sequence from the NCoA2 protein (sequence KENALLRYLLDKD), containing the signature motif LXXLL for binding to an NR (14). The structural models were prepared for molecular dynamics (MD) simulations using the AMBER21 package (21). Overall, four different systems were prepared: FXR wildtype in the active state (hereafter termed “active WT”), FXR T296I variant in the active state (“active T296I”), FXR wildtype in the inactive state (“inactive WT”), and FXR T296I variant in the inactive state (“inactive T296I”). Postprocessing and analysis of the MD trajectories were performed with CPPTRAJ (22) implemented in AmberTools21 (21). For further details, please see the SI Methods.

Statistical Analysis

Significance tests were performed using the Mann-Whitney U test or Student’s t-test if not indicated otherwise. The indicated significance levels are n.s. (not significant, $p > 0.05$), *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$, ****: $p \leq 0.0001$.

Results

The T296I variant is located within the LBD of FXR

The ligand binding domain (LBD) of FXR is critical in regulating the protein's activity. Residue 296 is located on helix 3 with its side chain facing toward the AF2 interaction surface formed partly by H12 (Fig. 1A and B). We thus hypothesized that variant T296I impacts FXR's ability to transition from the inactive to the active state. Accordingly, we investigated the effect of variant T296I in MD simulations starting from one of the four configurations: FXR WT in the active state ("active WT"), FXR T296I variant in the active state ("active T296I"), FXR WT in the inactive state ("inactive WT"), and FXR T296I variant in the inactive state ("inactive T296I") (Fig. 1C). All systems contained the LBD of FXR, the agonist chenodeoxycholic acid (CDCA) (23, 24), and a short peptide sequence of the nuclear receptor coactivator 2 (NCoA2). The inactive state was created from the active state through repositioning of the loop region between H11 and H12 such that H12 pointed away from the LBD core and had a distance > 45Å to it (distance in the active state 16Å). This setup allowed us to study if the substitution impacts the active state and/or the transition from the inactive to the active state.

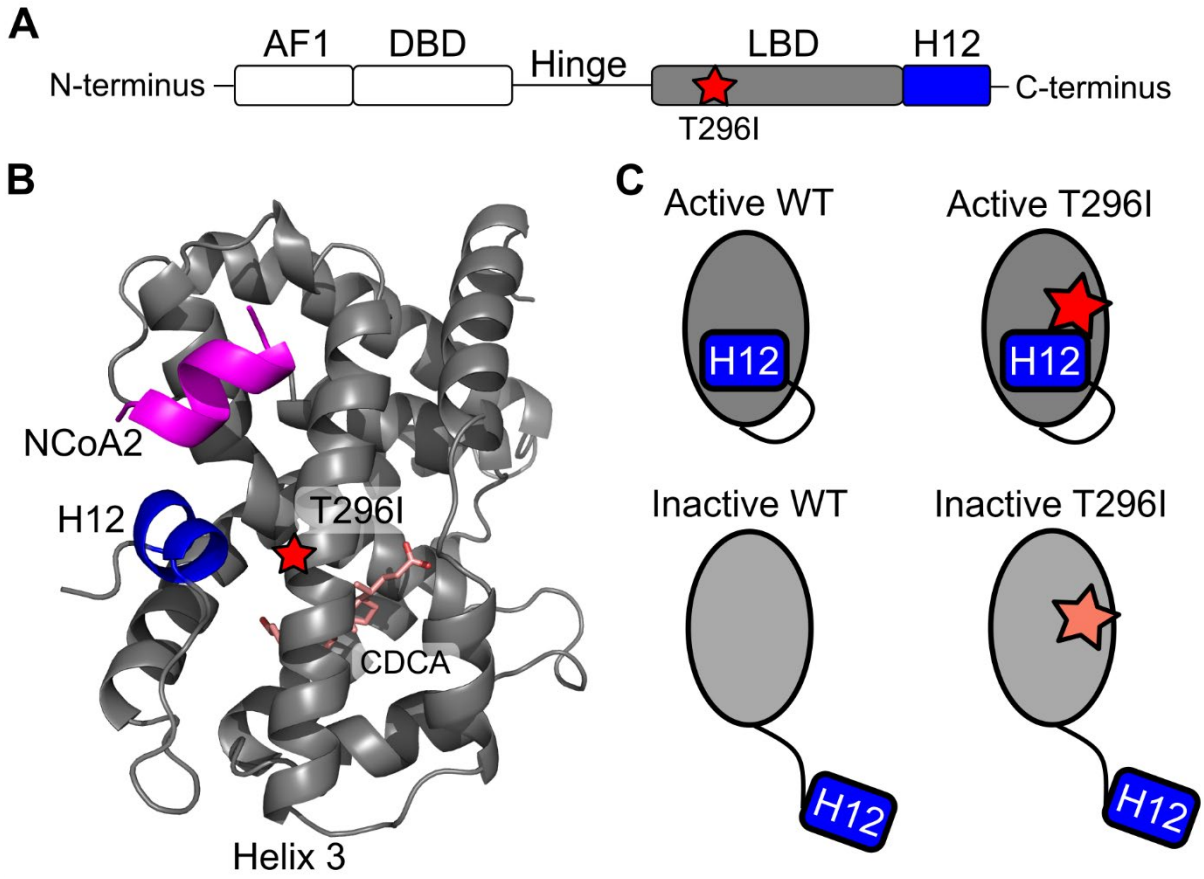

Fig. 1: Overview of the protein structure of FXR and the variant site within the LBD. A) Schematic of the domain arrangement of the FXR protein. The N-terminal activation function 1 (AF1) motif is followed by the DNA-binding domain (DBD), which is connected via a flexible hinge region to the LBD and the C-terminal H12. The variant T296I (red star) is located within the LBD. B) Protein structure of the LBD of FXR. The protein systems were modeled based on the crystal structure of agonist-bound FXR LBD (13) and used for MD simulations, containing additionally CDCA as ligand (pink, shown as sticks) and a short peptide of NCoA2 (magenta). H12 is highlighted in blue and H3, containing the variant site T296I (red star), is labeled. C) Overview of the four systems used as input to MD simulations to study the variant's impact on the active and the inactive state of the LBD. To differentiate between the different systems, we have consistently used the following color scheme: active WT in darker grey, inactive WT in lighter grey (corresponding to the depicted color of the LBD), active T296I in red and inactive T296I in faded red (corresponding to the depicted color of the star indicating the variant position).

T296I decreases the transcriptional activity of FXR

To determine the consequences of the FXR missense variant T296I on expression, subcellular localization and target gene induction, human FXRα1 and FXRα2 were cloned from human liver and co-transfected with RXRα into HEK293 cells. Both wildtype (WT) as well as the missense variant were detected within the nucleus of transfected cells (Fig. 2A). Furthermore, protein amounts as determined by western blotting were similar in WT and T296I transfected

cells (Fig. 2B). For functional analysis, we used a luciferase expression vector containing the BSEP promoter sequence ($BSEP^{prom}$-Luc), which was co-transfected with several combinations of RXRα and either FXRα1, FXRα2, $FXRα1^{T296I}$, or $FXRα2^{T296I}$ and subsequently stimulated with an FXR ligand (OCA, 10μM) and an RXR ligand (9-cis-RA, 1μM). The highest BSEP transactivation was observed when both RXR and WT FXRα1/2 were co-transfected, represented as 100% luciferase activity (Fig. 2C). Transfection of FXRα1 or FXRα2 alone resulted in luciferase activity of 19.24% and 34.45% respectively, in comparison to co-transfection with both RXRα and either WT FXRα1/2. However, when the $FXRα1/2^{T296I}$ variant was transfected alone, there was a significant decrease in luciferase activity to 11.72% and 4.89%, respectively, when compared to either WT FXRα1/2. Similarly, the co-transfection of RXRα with the $FXRα1/2^{T296I}$ variant resulted in a significant reduction in luciferase activity, with decreases to 26.06% and to 13.14%, respectively, in comparison to the co-transfection of RXRα and either WT FXRα1/2. We obtained similar results by using the SHP promoter sequence in the same luciferase expression vector (SHP^{prom}-Luc) and under the same experimental conditions (Fig. S2). In summary, even though the subcellular localization and protein expression levels were unaffected, the presence of the T296I missense variant resulted in a substantial decrease in FXR target gene transactivation *in vitro*.

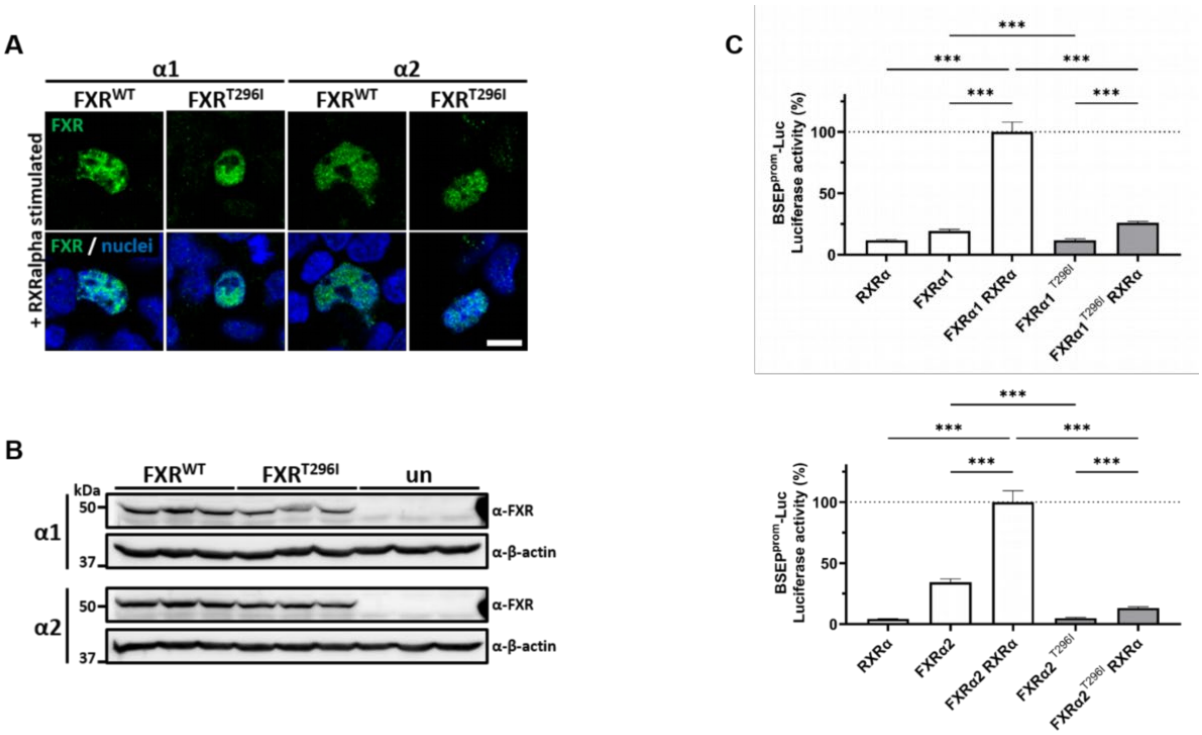

Fig. 2: FXR T296I reduces transcriptional activity of the BSEP promoter in transfected HEK293 cells. A) HEK293 transiently co-transfected with RXRα and either FXRα1, FXRα2, FXR$\alpha 1^{T296I}$, or FXR$\alpha 2^{T296I}$ showed correct nuclear localization of the wildtype (WT) and mutant protein. Bar = 10µm. B) Western blot analysis revealed similar protein amounts in cells transfected with the different FXR cDNA constructs, un=untransfected controls. C) Analysis of luciferase enzymatic activity after transfection of HEK293 cells with a luciferase reporter gene downstream of the BSEP promoter ($BSEP^{prom}$-Luc) as well as different combinations of RXRα and either FXRα1, FXRα2, FXR$\alpha 1^{T296I}$, or FXR$\alpha 2^{T296I}$, as indicated on the x-axis. The plasmid pRL-TK was included in each transfection for normalization. Cells were stimulated with the FXR ligand OCA (10µM) and the RXR ligand 9-cis-RA (1µM). Values were obtained from three independent experiments, in which each condition was tested in duplicates. Values on the y-axis represent the mean and SD, expressed as % luciferase activity. The asterisks indicate a significant difference analyzed by a two-tailed Student t-test, *** = $p \leq 0.001$.

FXR T296I leads to an increased distance between H12 and the substitution site, indicative of a less favorable active state

Using MD, we investigated the molecular mechanisms underlying the decreased activity of the T296I variant. All simulation systems showed minor structural variability with respect to the binding of CDCA and NCoA2 and the FXR LBD structure up to and including H11 (Fig. S1).

Based on the crystal structure of the FXR LBD (13), representing the agonist-bound active state, the WT T296 likely interacts with T466 preceding H12 (Fig. 3A). Accordingly, we measured the distance between residue T296 and T466 during MD simulations and compared it to the reference distance in the crystal structure. For the active states, the variant showed an increase in the distance (Fig. 3B). We determined the frequency of occurrence when the reference cutoff distance was reached (Fig. 3C, Table S1). Across the 15 replicas, the active WT system was found in approx. 27% of the time in the active state according to distances below the reference distance. The frequency of occurrence was significantly lower for active T296I (0.40%). The inactive systems showed initially high distance values, as expected. Inactive WT reached below the reference distance in 6 out of 15 replicas and often stayed within this active state for the remainder of the simulation time, indicating that the active state is the preferred one under the simulation conditions (Fig. S3, Table S1). Inactive T296I only reached the reference distance in one replica (Fig. S3, Table S1) and, accordingly, the frequency of reaching the reference value was significantly reduced in the inactive T296I system compared to the inactive WT system (Fig. 3C).

Overall, we observed an increase in the distance between residue T296 and T466 for the T296I variant in the active state. In line, for systems started either from the active or inactive state, T296I led to a significant decrease in the frequency of occurrence in reaching the reference distance compared to the WT. This data indicates that the active state is destabilized in the variant.

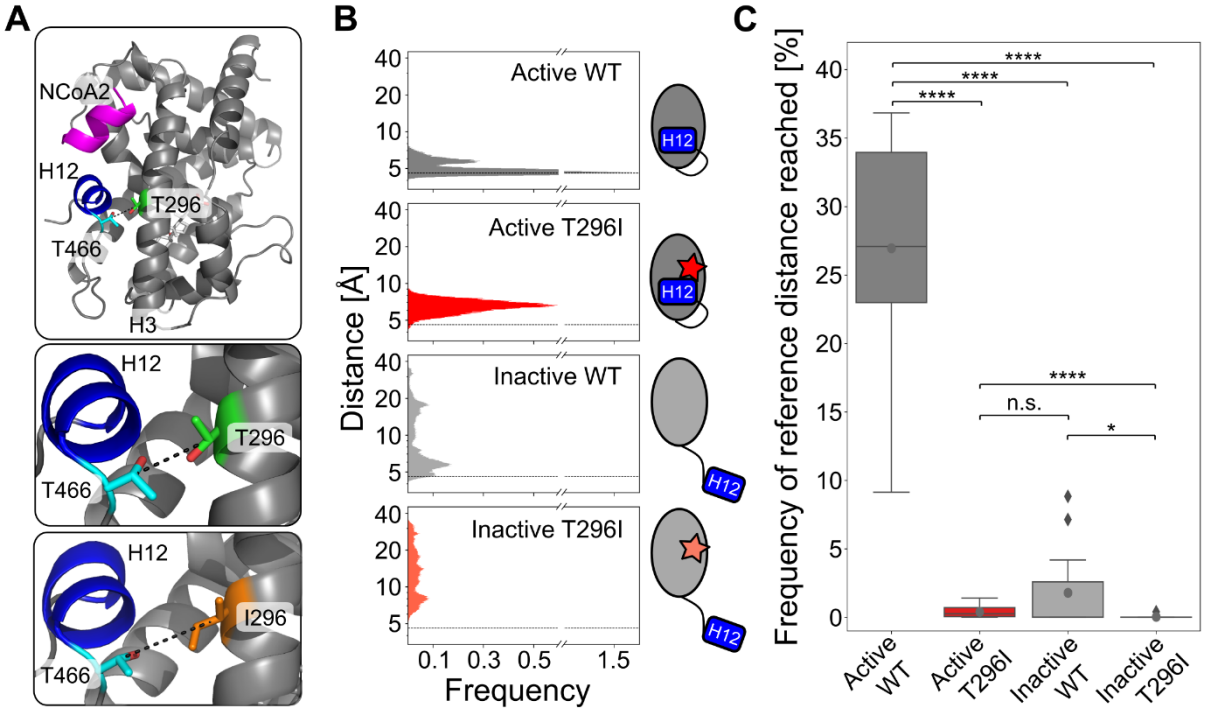

Fig. 3: Distance analysis between residues T296 and T466 over MD simulation time. A) Depiction of the distance measured within the LBD of FXR between the C_β atoms of residue T296 and T466. The mean distance is increased in the active T296I (6.6Å) compared to the active WT (5.0Å). B) Histograms of measured distances over all 15 MD runs (see Fig. S3) for each analyzed state. The reference distance (4.6Å) as measured in the agonist-bound crystal structure (13) is depicted as dashed lines. C) Frequency of occurrence that the respective system is in the active state. For each replica (Fig. S3), the percentage of reaching a distance below the reference distance (4.6Å) is depicted as a boxplot. Individual values are shown in Table S1. Boxes depict the quartiles of the data with the median (straight black lines) and mean (grey dots) indicated; the whiskers indicate the minimum and the maximal values, outlier points are depicted as rhombus. Differences in the mean values were statistically evaluated using a two-sided Mann-Whitney U test (N = 15, n.s.: not significant; *: $p \leq 0.05$, ****: $p \leq 0.0001$).

The correct positioning of H12 in the active state is reduced in the T296I variant

As the correct positioning of H12 within NRs is crucial for activity (25-28), we visually analyzed the simulation trajectories of the inactive systems and exemplarily show the conformational change from the inactive to the active state for one out of several MD replicas of the inactive WT system showing this transitioning (Fig. 4A and Movie S1). For the inactive T296I, we show the MD replica where H12 positioning was closest to the conformation in the active state (Fig. 4B and Movie S2). Further, we analyzed the impact of T296I on the positioning of H12 using the root mean square deviation (RMSD) of residues 466 to 473 (H12 and preceding T466) over the MD simulation time (Fig. 4C and D). As a reference state, we employed the active

conformation based on the crystal structure of the agonist-bound FXR LBD (13). For the active WT, a skewed Gaussian function was fitted to the RMSD histogram, yielding a mean of 1.9Å (Fig. S4 and Fig. 4C). Over the entire MD simulations, more than half of the time (~56%, Table S2 and Fig. 4D) the active WT had an RMSD below this mean value, which we used as a further reference to indicate reaching the active state. Comparing the active systems revealed a significant shift in the distributions, with a larger mean of active T296I indicating a higher deviation from the active positioning (active WT mean: 1.9Å, active T296I mean: 2.6Å, p = 0.0022, two-sided t-test, Fig. S4). The frequency of reaching the active state was significantly lower for active T296I, inactive WT, and inactive T296I systems compared to active WT (Fig. 4D). While the inactive WT reached the reference RMSD value in four out of 15 replicas, the inactive T296I reached it in one replica (Fig. S5). Furthermore, while the inactive T296I did not stay in the active state long (frequency: 0.01% in replica no. 6), the inactive WT – once reaching the active state – showed often prolonged persistence times (frequency: 15.40% in replica no. 2, 0.26% in replica no. 6, 10.17% in replica no. 8, 14.29% in replica no. 13) (Fig. S5 and Table S2). The comparison between inactive WT and inactive T296I indicated a similar trend as observed for the distance analysis but did not reach the significance level (Fig. 4D).

Overall, H12 positioning is significantly structurally deviating with respect to the reference active state for all systems compared to the active WT. While the active T296I could reach the reference cutoff, it did so for a significantly decreased amount of time compared to the active WT, again indicating that the active conformation is less favorable in the variant. Although inactive WT and inactive T296I could both reach the cutoff, the inactive WT reached it more frequently and for a longer time. However, the differences to inactive T296I are not significant. Our data indicate that unbiased MD simulations on the μs-scale can sample the transition from the inactive to the active state (see Table 1) and that this conformational change is less frequent in the variant.

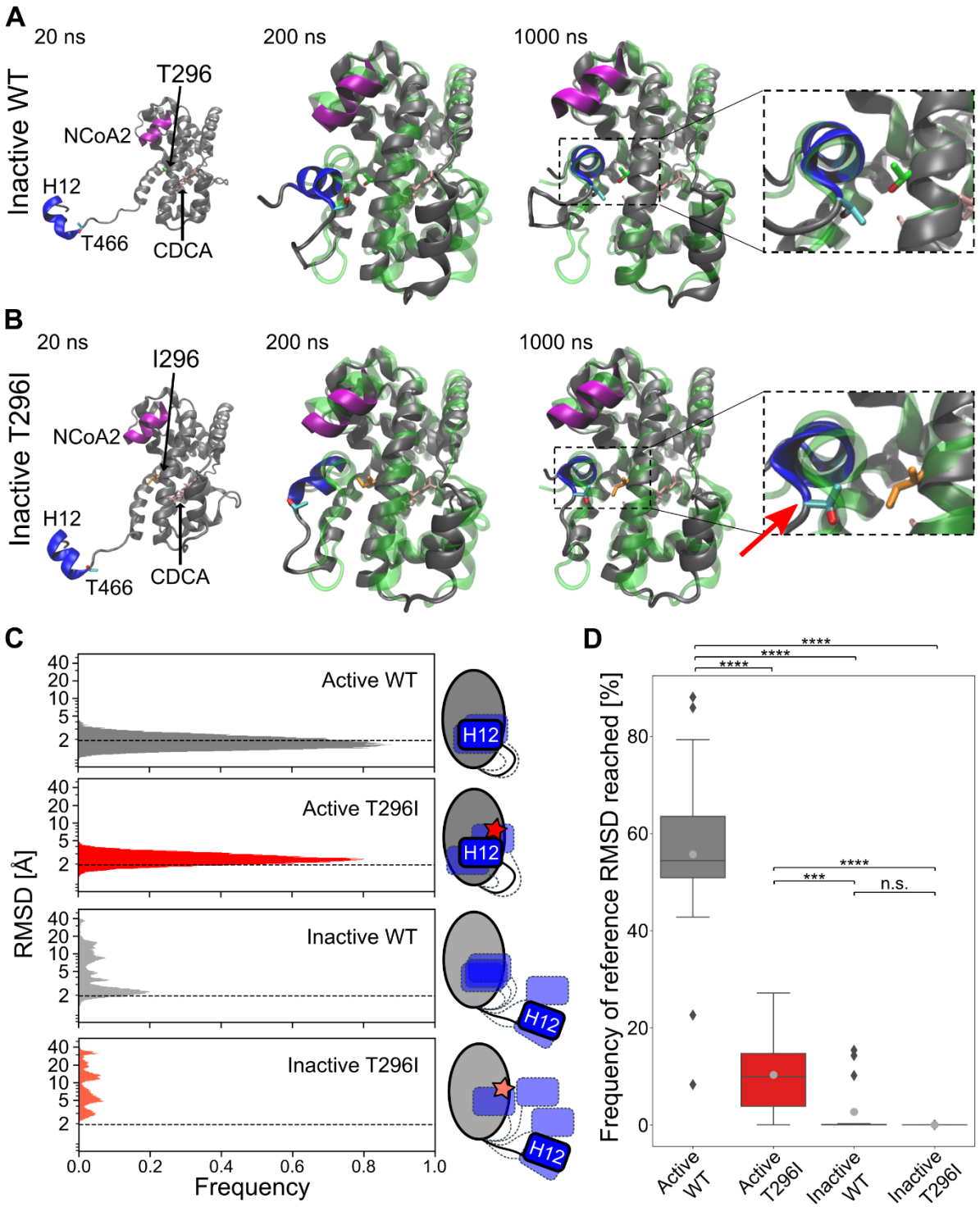

Fig. 4: Positioning of H12 during MD simulations. A) Conformational transitioning of the inactive WT (replica no. 2, Fig. S3 and S5) over the MD simulation time. The initial active state (based on the crystal structure of agonist-bound FXR LBD (13) is depicted as a green, translucent cartoon structure. B) Transitioning of the inactive T296I (replica no. 6, Fig. S3 and S5) over the MD simulation time. The initial active state is depicted as a green, translucent cartoon structure. Important residues (variant site 296 [green or orange, shown as sticks], T466 [cyan, shown as sticks], H12 [blue], NCoA2 [magenta], and bound ligand CDCA [pink, shown as sticks]) are additionally depicted in A and B. After 1000ns, H12 of the inactive WT almost perfectly overlayed with that of the active structure, while H12 of the inactive T296I showed

structural deviations (red arrow). C) Histograms of RMSD values of H12 and the preceding T466 over all 15 MD replicas for each analyzed state. Data for each MD replica are depicted in Fig. S5. A skewed Gaussian function was fitted to the distribution of active WT (Fig. S4), and the obtained mean (1.9Å) was used as a reference cutoff (dashed lines). D) Frequency of occurrence a system spends in the active state, i.e., when the reference cutoff is reached. For each replica (Fig. S5), the frequency of occurrence was determined and depicted within boxplots. Individual values are shown in Table S2. Boxes depict the quartiles of the data with the median (straight black lines) and mean (grey dots) indicated; the whiskers depict the minimum and the maximum values, outlier points are depicted as rhombus. Differences in the mean values were statistically evaluated using a two-sided Mann-Whitney U test (N = 15, n.s.: not significant; *: ***: $p \leq 0.001$, ****: $p \leq 0.0001$).

Table 1: Overview of the MD simulation results.

	Distance criteria	RMSD criteria
Inactive WT reaching active state	6 out of 15 runs	4 out of 15 runs
Inactive T296I reaching active state	1 out of 15 runs	1 out of 15 runs
Significance between WT and T296I[a]	* (p = 0.016)	n.s. (p = 0.076)

[a] Using one-sided *t*-test.

The T296I variant is associated with reduced expression of FXR target genes

To investigate the mRNA expression of FXR and two of its targets (BSEP, SHP) in the PFIC5 patient carrying the T296I variant, we performed qPCR analysis using FFPE samples from the patient´s liver taken at the time of transplantation. Additionally, FFPE-liver samples from two cirrhotic adult patients and a healthy adult control were included in the analysis. To address the putatively low RNA integrity after isolation from FFPE samples, we employed TaqMan Gene Expression Assays targeting different regions of the FXR and BSEP transcripts. While FXR mRNA expression was similar between the patient and control and the two cirrhosis livers (Fig. 5A), BSEP mRNA and SHP mRNA expression was strongly reduced in comparison to the healthy control but also the two cirrhotic liver samples (reduction to 3.03, 1.82 and 10.91% of healthy control for BSEP Taq1, BSEP Taq2, and SHP, respectively) (Fig. 5B). These

findings further demonstrate that our *in vitro* and *in silico* data authentically reflect the impaired transcriptional target gene activation by the FXR T296I variant.

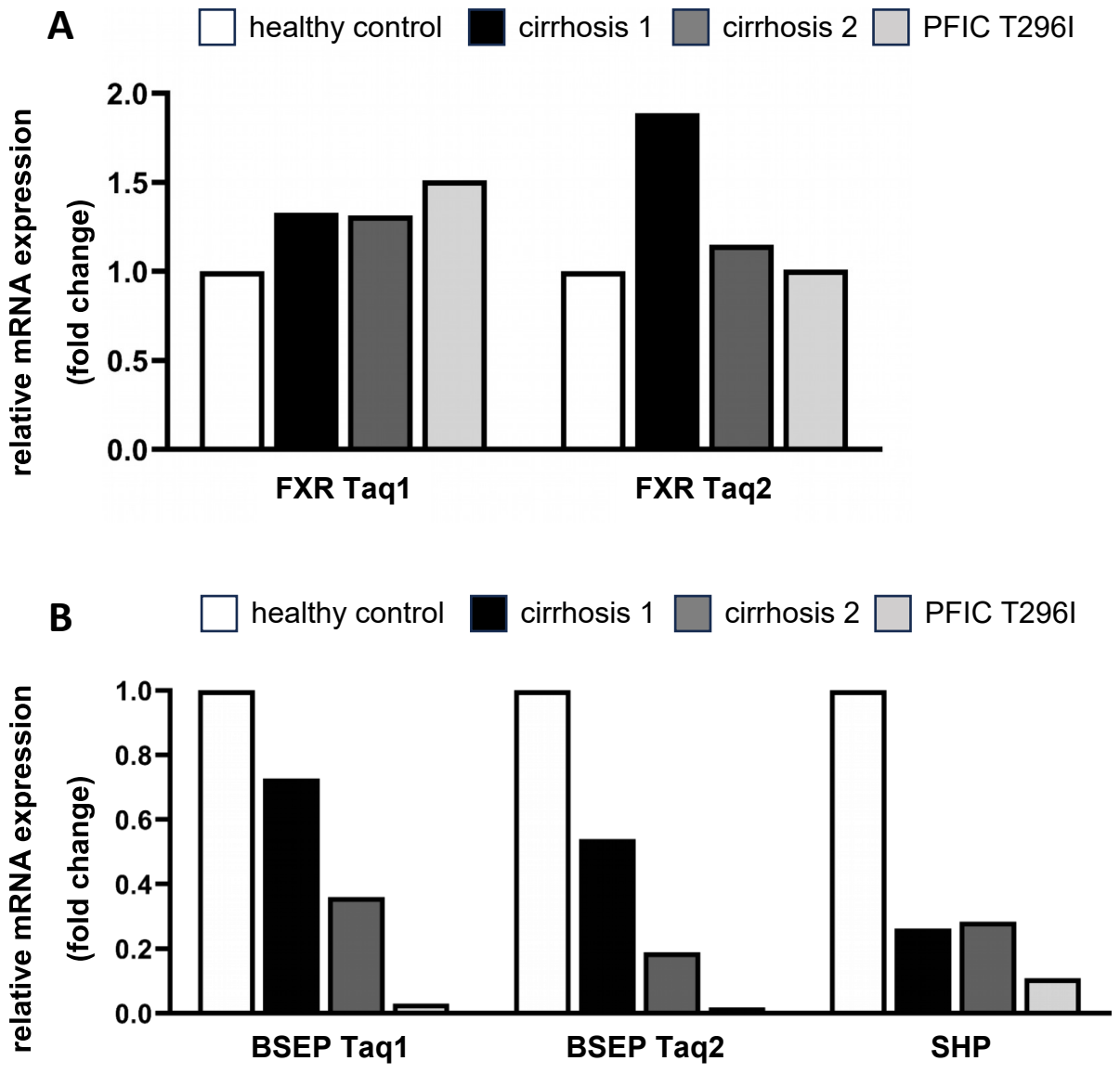

Fig. 5: FXR, BSEP, and SHP expression in the patient's liver tissue. A) Relative mRNA expression of FXR using two different TaqMan probes (Taq1 and Taq2) showed similar levels in the liver of a healthy control (white bars), two samples from cirrhotic livers (black bar and dark grey bar) as well as the patient (light grey bar). B) FXR target gene expression of BSEP and SHP were lower in the two cirrhotic livers as compared to the healthy control but were further reduced in the sample from the T296I PFIC patient.

Discussion

In this study, we combined *in vitro* experiments with computational studies to analyze the impact of the PFIC5-associated NR1H4 T296I variant on FXR protein expression, subcellular localization, and function. While the introduction of the missense variant into the α1 or α2 FXR isoform did not affect protein expression and nuclear localization *in vitro*, it significantly reduced activation of the FXR target genes BSEP and SHP. A strong reduction of BSEP and SHP was also observed in the patient's liver at the time of transplantation. Using our computational approach, we elucidated a detailed mechanism for the effect of the variant on the conformational transition of the LBD from the inactive to the active state. The variant showed a significantly reduced tendency to reach the active state, which can explain the *in vitro*-identified decreased target gene expression and thus the PFIC phenotype of the patient (9).

We describe conformational changes from an inactive to the known active state of the FXR LBD in unbiased MD simulations. To drive the system towards the active state, we used a coactivator peptide and the most potent *in vivo* agonist CDCA (29) within the MD systems as both coactivator peptide and ligand binding have been shown to induce and stabilize the active state *in vitro* (13). Depending on the analysis, the inactive WT system reached the active state in 27% or 40% (4/15 replicas for H12 RMSD analysis and 6/15 replicas for distance analysis, respectively) of the simulations (Table 1). A dynamic movement from the inactive to the active state (and potentially reverse, at least in ligand-free states) may occur in the nanosecond time scale as indicated by time-resolved fluorescence anisotropy decay studies on PPARγ (27). Chrisman et al. showed conformational changes of the H12 within the NR PPARγ towards an almost-active state within the µs to ms time scale range in unbiased MD simulations (30). Within our MD simulations, the LBD of FXR might sample conformational spaces usually not available due to sterical hindrances either by the not considered parts of FXR (DNA binding domain and linker sequences) or due to hetero-dimer binding partners. Thus, it is not surprising that in several replicas, H12 did not move into the active position within the 1µs of simulation time due to being trapped in other energy minimas. Still, fast transitioning from the inactive to the active state was observable in some replicas. Comparing the difference of inactive WT and

inactive T296I systems in reaching the active state, we observed a decrease by a factor of 3.8 to 5.7 (inactive WT: 26.7% or 40%; inactive T296I: 7%). This is in good agreement with the transcriptional activity reduction of T296I compared to WT as shown in the luciferase assay for BSEP as well as the patient's liver tissue (Fig. 2, Fig. 5).

Overall, the variant T296I, while also impacting the active state, likely exerts its negative impact on protein activity due to a change in the structural dynamics of the inactive-to-active state transition. Our results indicate that the T296I protein does not reach the active state fully and less frequently compared to the WT protein.

Furthermore, from the analysis of the MD simulations, insights into the activation mechanism of the LBD were gained. The presence of the ligand and co-activation peptide allows FXR to switch into an active state and stably stay within this state. This is in line with previous NMR or MD studies in RORγ (11), PPARγ (30, 31), and FXR or FXR/RXR heterodimers (16-18). By contrast, the transition from inactive to active FXR has previously not been observed in MD simulations. Our setup of the MD simulations can be useful to predict the impact of other missense variants on FXR function and potentially strengthen studies on FXR targeting, enabling detailed evaluations of the molecular mechanism of drugs based on their impact on the activation transitioning.

References

1. Bull LN, Thompson RJ. Progressive Familial Intrahepatic Cholestasis. Clin Liver Dis 2018;22:657-669.
2. Felzen A, Verkade HJ. The spectrum of Progressive Familial Intrahepatic Cholestasis diseases: Update on pathophysiology and emerging treatments. Eur J Med Genet 2021;64:104317.
3. Sinal CJ, Tohkin M, Miyata M, Ward JM, Lambert G, Gonzalez FJ. Targeted disruption of the nuclear receptor FXR/BAR impairs bile acid and lipid homeostasis. Cell 2000;102:731-744.
4. Cariello M, Piccinin E, Garcia-Irigoyen O, Sabba C, Moschetta A. Nuclear receptor FXR, bile acids and liver damage: Introducing the progressive familial intrahepatic cholestasis with FXR mutations. Biochim Biophys Acta Mol Basis Dis 2018;1864:1308-1318.
5. Gomez-Ospina N, Potter CJ, Xiao R, Manickam K, Kim MS, Kim KH, Shneider BL, et al. Mutations in the nuclear bile acid receptor FXR cause progressive familial intrahepatic cholestasis. Nat Commun 2016;7:10713.
6. Himes RW, Mojarrad M, Eslahi A, Finegold MJ, Maroofian R, Moore DD. NR1H4-related Progressive Familial Intrahepatic Cholestasis 5: Further Evidence for Rapidly Progressive Liver Failure. J Pediatr Gastroenterol Nutr 2020;70:e111-e113.
7. Mehta S, Kumar K, Bhardwaj R, Malhotra S, Goyal N, Sibal A. Progressive Familial Intrahepatic Cholestasis: A Study in Children From a Liver Transplant Center in India. J Clin Exp Hepatol 2022;12:454-460.
8. Czubkowski P, Thompson RJ, Jankowska I, Knisely AS, Finegold M, Parsons P, Cielecka-Kuszyk J, et al. Progressive familial intrahepatic cholestasis - farnesoid X receptor deficiency due to NR1H4 mutation: A case report. World J Clin Cases 2021;9:3631-3636.
9. Pfister ED, Dröge C, Liebe R, Stalke A, Buhl N, Ballauff A, Cantz T, et al. Extrahepatic manifestations of progressive familial intrahepatic cholestasis syndromes: presentation of a case series and literature review. Liver International 2022.
10. Hollingsworth SA, Dror RO. Molecular Dynamics Simulation for All. Neuron 2018;99:1129-1143.
11. Saen-Oon S, Lozoya E, Segarra V, Guallar V, Soliva R. Atomistic simulations shed new light on the activation mechanisms of RORγ and classify it as Type III nuclear hormone receptor regarding ligand-binding paths. Scientific Reports 2019;9:17249.
12. Chrisman IM, Nemetchek MD, de Vera IMS, Shang J, Heidari Z, Long Y, Reyes-Caballero H, et al. Defining a conformational ensemble that directs activation of PPARgamma. Nat Commun 2018;9:1794.
13. Merk D, Sreeramulu S, Kudlinzki D, Saxena K, Linhard V, Gande SL, Hiller F, et al. Molecular tuning of farnesoid X receptor partial agonism. Nat Commun 2019;10:2915.
14. Heery DM, Kalkhoven E, Hoare S, Parker MG. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature 1997;387:733-736.
15. Xu HE, Stanley TB, Montana VG, Lambert MH, Shearer BG, Cobb JE, McKee DD, et al. Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPARalpha. Nature 2002;415:813-817.
16. Kumari A, Mittal L, Srivastava M, Pathak DP, Asthana S. Conformational Characterization of the Co-Activator Binding Site Revealed the Mechanism to Achieve the Bioactive State of FXR. Front Mol Biosci 2021;8:658312.
17. Kumari A, Mittal L, Srivastava M, Pathak DP, Asthana S. Deciphering the Structural Determinants Critical in Attaining the FXR Partial Agonism. J Phys Chem B 2023;127:465-485.

18. Díaz-Holguín A, Rashidian A, Pijnenburg D, Monteiro Ferreira G, Stefela A, Kaspar M, Kudova E, et al. When Two Become One: Conformational Changes in FXR/RXR Heterodimers Bound to Steroidal Antagonists. ChemMedChem 2023;18:e202200556.
19. Lee HK, Lee YK, Park SH, Kim YS, Park SH, Lee JW, Kwon HB, et al. Structure and expression of the orphan nuclear receptor SHP gene. J Biol Chem 1998;273:14398-14402.
20. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 2018;46:W296-W303.
21. Case DA, Aktulga HM, Belfon K, Ben-Shalom IY, Borzell SR, Cerutti DS, Cheatham TE, 3rd, et al. Amber. In. University of California, San Francisco; 2021.
22. Roe DR, Cheatham TE, III. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. Journal of Chemical Theory and Computation 2013;9:3084-3095.
23. Makishima M, Okamoto AY, Repa JJ, Tu H, Learned RM, Luk A, Hull MV, et al. Identification of a nuclear receptor for bile acids. Science 1999;284:1362-1365.
24. Parks DJ, Blanchard SG, Bledsoe RK, Chandra G, Consler TG, Kliewer SA, Stimmel JB, et al. Bile acids: natural ligands for an orphan nuclear receptor. Science 1999;284:1365-1368.
25. Mi LZ, Devarakonda S, Harp JM, Han Q, Pellicciari R, Willson TM, Khorasanizadeh S, et al. Structural basis for bile acid binding and activation of the nuclear receptor FXR. Mol Cell 2003;11:1093-1100.
26. Renaud JP, Rochel N, Ruff M, Vivat V, Chambon P, Gronemeyer H, Moras D. Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. Nature 1995;378:681-689.
27. Kallenberger BC, Love JD, Chatterjee VK, Schwabe JW. A dynamic mechanism of nuclear receptor activation and its perturbation in a human disease. Nat Struct Biol 2003;10:136-140.
28. Nolte RT, Wisely GB, Westin S, Cobb JE, Lambert MH, Kurokawa R, Rosenfeld MG, et al. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-γ. Nature 1998;395:137-143.
29. Wang H, Chen J, Hollister K, Sowers LC, Forman BM. Endogenous bile acids are ligands for the nuclear receptor FXR/BAR. Mol Cell 1999;3:543-553.
30. Chrisman IM, Nemetchek MD, de Vera IMS, Shang J, Heidari Z, Long Y, Reyes-Caballero H, et al. Defining a conformational ensemble that directs activation of PPARγ. Nature Communications 2018;9:1794.
31. Heidari Z, Chrisman IM, Nemetchek MD, Novick SJ, Blayo A-L, Patton T, Mendes DE, et al. Definition of functionally and structurally distinct repressive states in the nuclear receptor PPARγ. Nature Communications 2019;10:5825.