scientific reports

OPEN

Molecular dynamics, docking and quantum calculations reveal conformational changes influenced by CYP271A amino acid mutations related to cerebrotendinous xanthomatosis

Yudibeth Sixto-López[1,2]✉, Humberto L. Mendoza-Figueroa[2], Bruno Landeros-Rivera[3], Alejandra Camacho-Molina[4] & José Correa-Basurto[2]✉

Cerebrotendinous xanthomatosis (CTX) is an autosomal recessive lipid disorder caused by a deficiency in CYP27A1, the first enzyme in the bile acid biosynthesis pathway. CYP27A1 catalyzes the 7α-hydroxylation of cholesterol, playing an important role in cholesterol homeostasis. CTX leads to progressive neurological dysfunction, including cognitive impairment, epilepsy, peripheral neuropathy, and movement disorders. Missense mutations in CYP27A1 disrupt its activity, particularly at the heme binding region and the adrenodoxin-binding site. This study examined the structural effects of seven-point mutations in CYP27A1 using molecular dynamic (MD) simulations. Both mutant and wild-type (WT) proteins were modeled to observe their structural behavior. Additionally, by combining MD simulations, docking, and quantum calculations cholesterol binding was studied in WT and mutant proteins. Results indicated that mutations altered cholesterol binding mode, preventing it from adopting the correct position in the catalytic site. The substrate access channel in mutants became wider, shallower, or closed. The interaction between the isopropyl group of cholesterol and the heme was found to be crucial for the hydroxylation capacity of CYP27A1, as this interaction was only present in the cholesterol-WT complex.

Keywords CYP27A1, Molecular dynamics simulations, Cerebrotendinous xanthomatosis, Molecular Docking

Cerebrotendinous xanthomatosis (CTX) is an autosomal recessive lipid storage disorder caused by point mutations in the CYP27A1 gene. These mutations cause a deficiency of mitochondrial sterol (25R)26-hydroxylase enzyme, also known as 27-sterol-α-hydroxylase (CYP27A1)[1,2]. CYP27A1 is the first enzyme that participates in the acid pathway of bile acid biosynthesis, catalyzing the oxidation of the terminal carbon of the side chain of cholesterol yielding (25R)26-hydroxycholesterol and subsequently to an cholestenoic acid (Fig. 1)[2]. In addition to cholesterol, CYP27A1 has another substates such as Vitamin D, cholestanol, and sterols[3,4]. Cholesterol oxidation generates several products with additional oxygen atoms, such as alcohols, ketones, epoxides, or carboxyl, also known as oxysterols. The presence of oxysterols in the human body is constant because of cholesterol biotransformation either by the reactivity against oxygen radicals or the catalytic activity of the CYP450 hydroxylases[5]. This alteration reduces the bile acids production (chenodeoxycholic acid and cholic acid). Consequently, some intermediate metabolites, such as bile alcohols, cholestanol, and other oxysterols accumulate and reach pathological levels

[1]Departamento de Química Farmacéutica y Orgánica, Facultad de Farmacia, Universidad de Granada, Campus de Cartuja s/n, 18071 Granada, Spain. [2]Laboratorio de Diseño y Desarrollo de Nuevos Fármacos e Innovación Biotecnológica (Laboratory for the Design and Development of New Drugs and Biotechnological Innovation), Escuela Superior de Medicina, Plan de San Luis y Díaz Mirón, Instituto Politécnico Nacional, Ciudad de México 11340, México. [3]Facultad de Química, Departamento de Química Inorgánica y Nuclear, Universidad Nacional Autónoma de México, Circuito exterior S/N, Ciudad Universitaria, Ciudad de México, México. [4]Consejo Mexicano de Genética A.C., Mexico City, México. ✉email: syudibeth@hotmail.com; corrjose@gmail.com

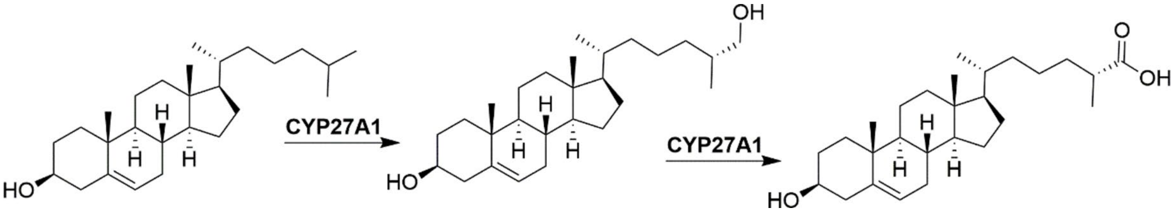

Fig. 1. The first two reactions catalyzed by CYP27A1 in the acidic bile acid pathway[2].

in tissues, mainly in the central nervous system, tendons, and cardiovascular system[6]. Progressive neurological dysfunction manifests as cognitive impairment, epilepsy, pyramidal and cerebellar signs, peripheral neuropathy, and/or movement disorders[7]. CTX is considered a rare disease; it has been observed more frequently among females than males (~ 55% vs. 45%). The prevalence of CTX in the US is estimated to be 3–5 per 100,000 people of European descent and the incidence among Americans is ~ 1:72,000 to 1:150,000. However, estimates of CTX incidence vary by region of the world[8]. CYP27A1 contains 531 amino acids, of which 33 correspond to the mitochondrial signal sequence and 498 to the mature protein. CYP27A1 is composed by the sterol-binding site (P384-P398), the adrenodoxin-binding (ADX) site (L351–V365), and as other CYPS it contains the conserved heme-binding site, oxygen binding domain, and the conserved ERR (glutamic acid-arginine-arginine) triad which is a region formed by two motifs, the K-helix motif (EXXR) and the electron transport channel (PERF), which are involved in the stabilization of the protein structure, ferredoxin binding and also in the transference of electrons to the heme iron[9–11].

Exhaustive molecular genetic analysis of the CYP27A1 gene identified more than 100 different mutations worldwide in CTX patients of different ethnicity, including missense, deletion, insertion, splice site, and nonsense mutations[12], and only 16 are missense mutations that could affect the expression of the CYP27A1 protein. It is predicted that 10 missense mutations have a relevant role in disrupting heme binding or the ADX site. The following point missense mutations lead to dysfunctional enzymes: R395C, R395S, R405Q, R474Q, R474W, and R479C[13]. Therefore, bioinformatic analysis of CTX-related natural variants may provide structural information on the active site and/or tertiary structure of sterol 27-hydroxylase[9,11].

Although, the clinical features, molecular, and biochemical basis of CTX are under study, there are limited descriptions of the structural and dynamic behavior of CYP27A1[3,14–16]. In the early 2000s, Sawada N. et al. conducted an experimental study focused on the structure-function relationship of CYP27A1 and some mutants (R127K, G145A, R405K, R474K) from CTX patients. The analysis revealed that these mutations disrupted the enzyme's structure and function, affecting the heme-binding and substrate-binding regions[3]. Further, Proser D. E., et al. modeled by homology modelling the structure of CYP27A1 and analyzed how specific mutations affect vitamin D hydroxylation through site-directed mutagenesis. They found that F248, T402, N403, S404, I514, V515, and L516 are crucial for substrate binding and the enzyme's function[14]. Mast N. et al., using molecular docking, established the binding mode of cholesterol. Combining this with site-directed mutagenesis, they found that mutations in W100, H103, T110, M301, V367, I481, and V482 affected substrate binding and enzyme activity in a substrate-dependent manner. Molecular docking determined that each substrate studied (cholesterol and 5β-cholestane-3α,7 α,12 α -triol) resides in different regions within the substrate-binding pocket and interacts with different residues. Particularly, it was found that V367 is important for positioning the cholesterol C26 methyl group[15]. In the most recent study, diverse bioinformatic tools were employed to determine the impact of mutations on CYP27A1 structure stability. Using docking studies the impact on cholecalciferol recognition was investigated. It was found that mutations led to changes in protein stability and binding affinities, suggesting potential impacts on enzyme function[16]. The state-of-the-art reports more structure-functional studies supported by docking and molecular modelling, leaving aside the dynamic behavior of the protein.

In the current study, we investigated the influences of amino acid mutations on structural mobility and molecular recognition of CYP27A1 homology models of 7 natural variants identified in CTX disease compared with the wild-type protein using computational tools, such as molecular dynamics, molecular docking, and quantum calculations.

Results and discussion

CTX is a rare disease caused by mutations in the CYP27A1 gene, leading to 27-sterol-α-hydroxylase malfunction. This enzyme catalyzes the initial oxidation step of cholesterol and other sterol intermediates in bile acid synthesis at the C-27 position[9,17]. In the early 2000s, several structure-functional studies were conducted, revealing the importance of certain point mutations in the heme-binding and substrate-binding regions, which affect the substrate binding and enzyme function, thereby establishing the molecular basis of the enzyme's function[3,14,15]. These studies combined experimental site-directed mutagenesis and enzyme activity determination with protein modelling and docking studies using substrates such as vitamin D, cholecalciferol, cholesterol and 5*β*-cholestane-3α,7 α,12 α -triol[3,14,15,18]. However, no exploration of the structural-dynamic behavior had been done. Therefore, in this research, we explored the structural and dynamic behavior of CYP27A1 and how it is affected by some clinically important point mutants. Additionally, we used molecular docking and quantum calculations to investigate the molecular interactions between cholesterol and CYP27A1 substrate, which can explain the impact on CTX.

3D models

No crystal structures are available for CYP27A1, but some 3D models have been published[14,15,18].

For this study, CYP27A1 was modeled by homology modelling using Swissmodel[19]. The 3D structure of rat mitochondrial CYP24A1 (PDB: 3K9Y) was used as a template[20] (Fig. 2). It shares 34.6% of sequence identity according to Swissmodel alignment and according to Bioedit[21]alignment it is shares 49.5% of similarity, which is suitable for reliable model prediction[22,23] (see full alignment at Supplementary material **Fig. S1**). Further, the template was structurally aligned with WT and mutants of CYP27A1, and no significant differences were found (RMSD between 0.26 and 0.651Å). Particularly, in CYPs it has been observed that the three-dimensional folding is well conserved between the members despite each one could only share 20% of the identity[10,24].

The model mainly consists of α-helices, with the heme group housed between the heme thiolate loop (where C476 is found), B-/B'loop, B'-C loop, helices (C and I) and flanked by a small β-sheets regions (β−3a) in the heme binding site (P435–F464) (Fig. 3). It shares the common P450 fold[14,25]. The typical substrate-access channel is observed, the conserved residues M247, F248, W268, Y271 and W275, which leads to the substrate-binding site[9,14,26] (Fig. 2A). The Fe atom of the heme group is coordinated with the well-conserved C476, and the heme propionic groups form hydrogen bonds with R158, W154, R474, R127, R405 and N403, which are in accordance with previous reports[3,14] (Fig. 2B and C). The conserved ERR triad is in K-helix (E392, R395) and loop between β−3b-L helix (S432, R433, S439, E440, S443, Q445, R448 and E392) (Fig. 2D). On the other hand, the putative alcohol-acid pair residues D338 and T339 (Fig. 2E), are found in the I-helix which in other CYPs form part of the oxygen-binding groove[20,27]. As can be seen, the generated model reliable reproduce the most important regions of the CYP27A1 and can be used for further studies.

3D model of CYP27A1 mutants

To study the effects of point mutations on CYP27A1, we modeled several CYP27A1 mutants. The first mutant, M1: G145E, is in the B′-C loop, which orients the substrate contact residues and is predicted to be pathogenic[3,14] (Fig. 2F). Other clinically relevant missense mutants (M2: R395C, M3: R395S, M4: R405Q, M5: R474Q, M6: R474W, and M7: R479C) were also modeled[13] (Fig. 2). As mentioned above, the models predicted by Swissmodel do not show significant structural differences based on structural alignment (RMSD = 0.260–0.651 Å). Important residues were compared with no major differences in the backbone or the side chains, except for the mutated residues highlighted in Fig. 2.

The R395C mutant (M2) drastically reduces the activity of the enzyme, likely due to its location in the conserved ERR triad region, which stabilizes the heme binding region (Fig. 2D). The ERR region acts as a folding motif and plays an important role in redox-partner binding[14,28]. The R395C mutation causes the loss of some interactions that may affect CYP27A1 folding and disrupt adrenodoxin binding[29]. Substitution by serine (M3) may also result in the loss of important interactions due Serine substitution can also result in the loss of

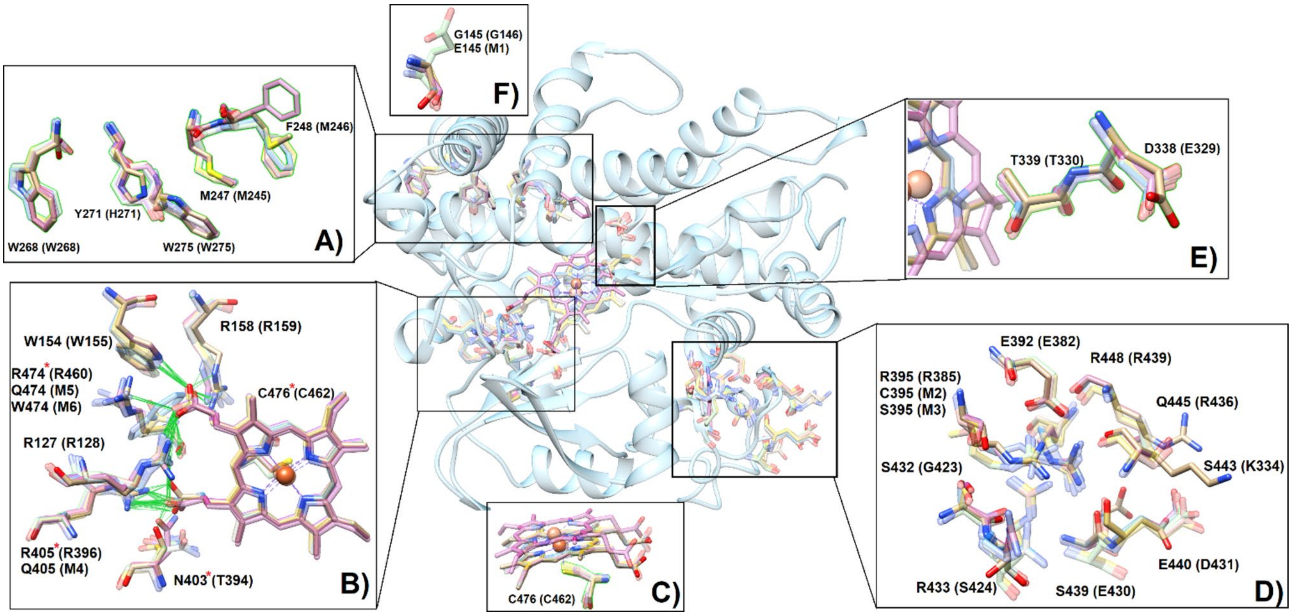

Fig. 2. CYP27A1 model generated by homology modeling. The crystal of CYP24A1 (PDB: 3K9Y) was used as a template. Important structural regions common in CYP450s were zoomed in by superimposing the template CYP24A1 residues in parentheses with respect to the mutants and WT. **(A)** Substrate access channel **I** (pw2a), **(B)** Hydrogen bond network of propionate-linked residues of the heme group, **(C)** C476 coordinated to Fe atom, **(D)** residues belonging to the ERR triad region, **(E)** residues of the alcohol-acid pair, and **(F)** Residue G145. Hydrogen bonds are represented as green lines. Residues belonging to the heme region are marked with *. The names of mutated residues are placed below the WT and the mutant is indicated between parentheses.

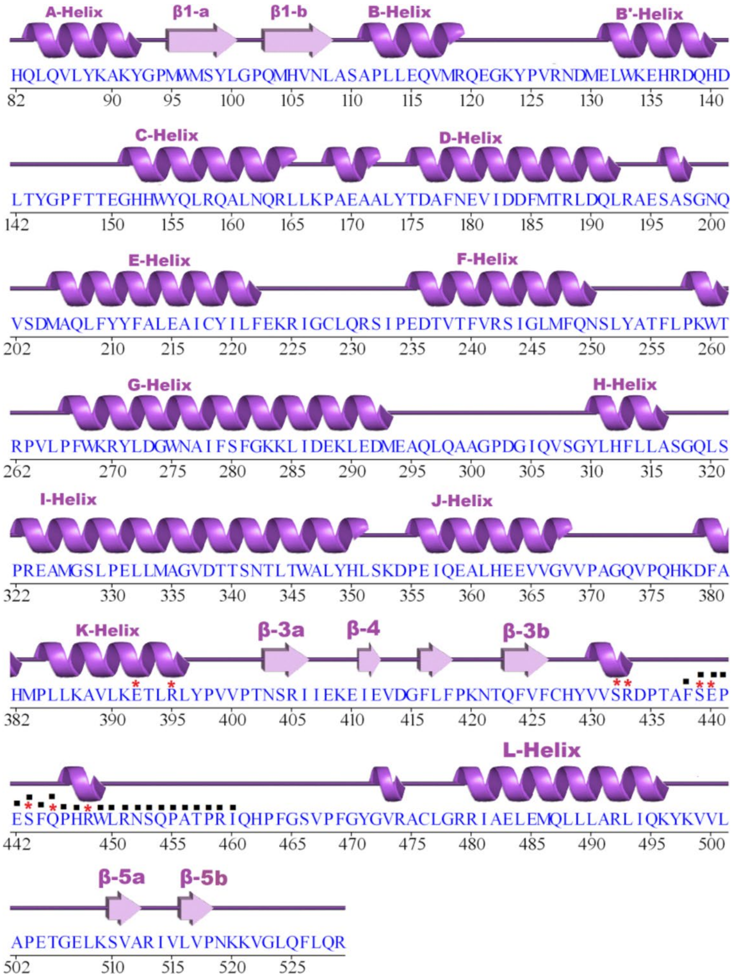

Fig. 3. The primary sequence of WT CYP27A1 aligned with the secondary structure of the WT model used in this study. The secondary structure notation was adapted from Prosser D., et al.[14], and created with PDBsum[82]. Residues that belong to the ERR region are marked with a red asterisk and those that belong to the meander region are marked with a black square[35,36].

important interactions due to the non-conservative mutation that eliminates the cationic character of arginine to the small polar side chain of serine (Fig. 2D). R395S mutation decreases the CYP27A1 activity, leading to CXT[30].

M4 (R405Q) is a mutation near the ADX binding site that could affect enzyme activity by disrupting a salt bridge between arginine and the heme propionate, stabilizing the heme group (Fig. 2B). A transient expression study showed that CYP27A1 lost its activity due to the R405Q mutation[31].

Mutations in R474 (M5: R474Q and M6: R474W) were studied (Fig. 2B). Both were reported as missense mutations and predicted to be pathogenic[32,33]. These mutations commonly occur in Japanese patients[33,34]. The R479C mutation (M7) significantly reduces the enzyme activity[28]. Once the models were modeled and validated (see Ramachandran plot in Supplementary material Fig. **S3-S10**), it was further submitted to MD simulation to refine the structure and study the dynamic and structural behavior of the WT and mutants CYP27A1.

Molecular dynamic simulations

Geometric parameters

To explore the dynamic and structural behavior of CYP27A1 protein, the apo WT and M1-M7 were submitted to 25-ns MD simulations using Amber 16. Root mean square deviation (RMSD) results show that most of the systems converged at approximately 10 ns. Mutations affect the structural behavior of the protein with all the mutants exhibiting more structural mobility compared to the WT. M1 (3.67 ± 0.16 Å), M5 (3.63 ± 0.39 Å), M6 (3.45 ± 0.19 Å), M4 (3.35 ± 0.23 Å), and M7 (3.24 ± 0.21 Å) showed more structural differences compared to the WT, while M2 (3.02 ± 0.32 Å) and M3 (2.90 ± 0.19 Å) showed similar dynamic behavior to the WT protein (2.88 ± 0.15 Å) (Fig. 4A). In terms of compactness, radius of gyration (Rg) suggests that M1-M2, and M4-M7 generated a slight expansion in the protein backbones compared to the WT protein; contrastingly M3 increases the protein compactness (Fig. 4B). Previous reports indicated that mutations in G145 (M1) affect the proper folding of the tertiary structure of the pocket where the heme group is housed; resulting in loss of enzyme activity due to conformational modifications, consistent with our findings for M1[3]. On the other hand, M4-

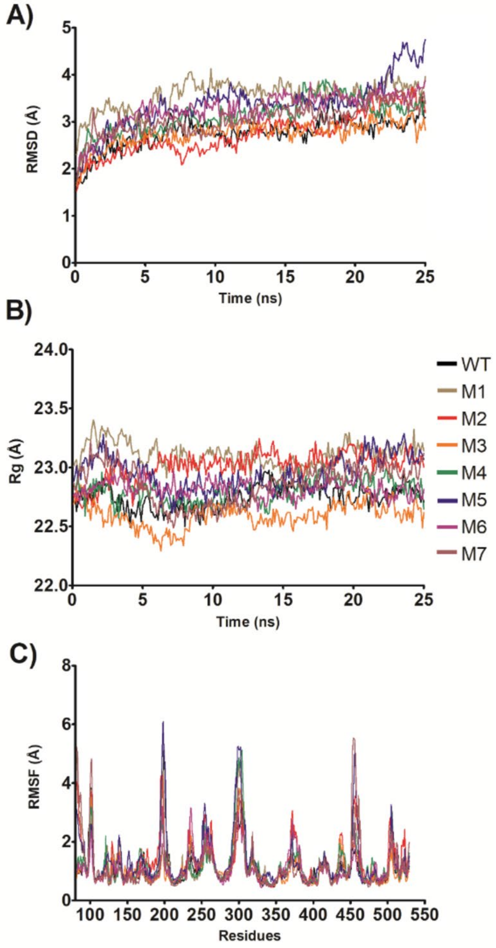

Fig. 4. Structural analysis of MD simulation of CYP27A1 wild type and mutants. **(A)** Root Mean Square Fluctuation (RMSF), **(B)** Root Mean Square Deviation (RMSD), and **(C)** radius of gyration (Rg).

M7 also showed structural differences compared to the WT. These findings agree with experimental reports, where mutations in R405 and R474 were associated with loss of enzyme activity due to the disruption of heme propionate coordination, while R479 disrupts heme incorporation[3,14].

The WT protein showed higher fluctuations (1.5–5.2 Å, Fig. 4C) in regions corresponding to loop portions located externally from the heme group, i.e., between the B-B'-helix (E121-L132), D-E helix (A194-Q207), G-H helix (L290-V307), J-K helix (P370-K380), and L-Helix-β–5a sheet (Y497-E507); the meander region (N452-R459), and loops between β–1a-β–1b sheets (S98-M104), and F-G helix (S251-V264), both forming the substrate access channel[14] (Fig. 5a). In CYP27A1 mutants, there were modifications in the protein backbone fluctuations compared to the WT (Fig. 4C). Increase fluctuations were observed in regions around the substrate access channel, A helix (H82-Y92) in M1, M2, M5, and M7, in β1-a and β1-b sheets (Y99-M104) in M7; and in loop regions between E-F-Helix (R231-T239) in M2-M6, G-H Helix (Q298-Q306) in M1, M3-5, and M7; and in loop between β sheet-3b-L helix (R451-S466) in M1-M2, M4-5, and M7, and loop between L-Helix-β–5a sheet (L501-S510) in M2, M4 and M5 (Fig. 5). In contrast, decrease fluctuations were observed in loops spanning β1-a and β1-b sheets (S98-M104) in M1-M6 which belong to the substrate access channel, and in the loop between D-E Helix (Q191-Q207) in M3-M4, M6, and M7, and L-Helix and β–5a sheet (P503-E507) in M3, M6, and M7.

Structural insights from most populated cluster conformation (MPCC)
A representative conformation of the most populated cluster from the MD simulations of each system were analyzed to understand structural changes due to the point mutations. When comparing the native conformation of the WT with the conformation obtained from the most populated cluster of the MD simulation, structural differences were observed as suggested by difference in the backbone with an RMSD = 1.149 Å (Fig. 5A). Differences were noted in loop regions between β–1a-β–1b sheets (S98-M104), B-B'-helix (E121-K123, D129-L132) and the F-G helix (S251-V264) (Table 1), which are part of the substrate access tunnel and the outer part of the binding cavity[14]. Additional differences were found in loops far from the catalytic site, such as those between the D-E helix (A194-207), G-H helix (L290-V307), J-K helix (P370-G372), β sheet-3B-L helix (F438-S443 and N452-R459), and L helix-β5a sheet (V500-E507).

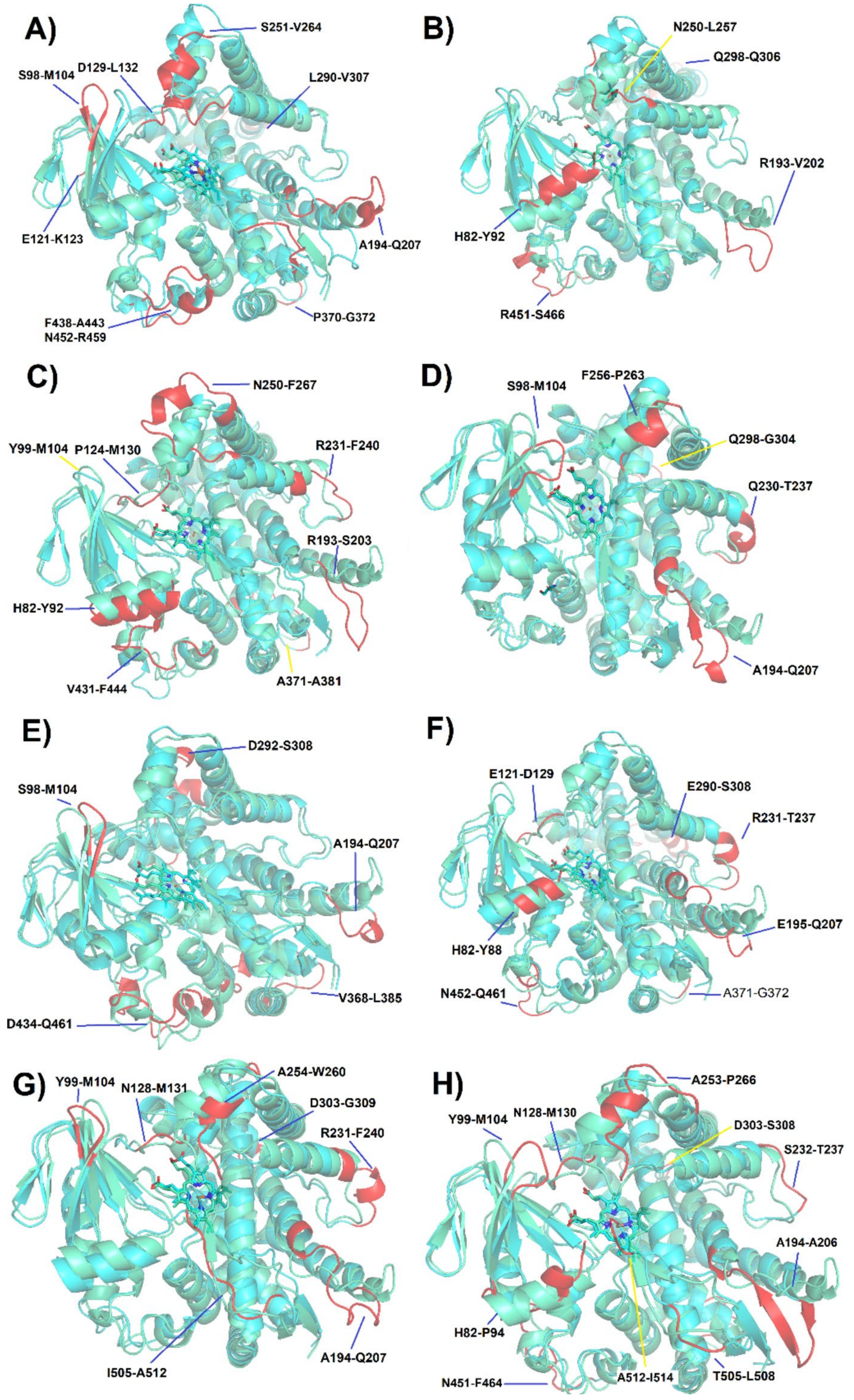

Fig. 5. Representative CYP27A1 conformations of the most populated cluster of WT and mutant proteins. **(A)** The native conformation (green) and the conformation retrieved from MPCC (cyan) are superimposed. The representative conformation of the WT of the most populated cluster (cyan) is superimposed with mutants **(B)** M1, **(C)** M2, **(D)** M3, **(E)** M4, **(F)** M5, **(G)** M6, and **(H)** M7, which are colored in cyan. Regions with major structural differences in the native conformation (A) or the conformation of the most populated cluster (B-H) are colored in red which is related to RMSF.

Model	Residues										
WT		S98-M104	E121-K123 D129-L132	A194-Q207		S251-V264	L290-V307	P370-G372	F438-A443 N452-R459	V500-E507	
M1	H82-Y92			R193-V202		N250-L257	Q298-Q306		R451-S466		
M2	H82-Y92	Y99-M104	P124-M130	R193-S203	R231-F240	N250-F267		A371-A381	V431-F444		
M3		S98-M104		A194-Q207	Q230-T237	F256-P263	Q298-G304				
M4		S98-M104		A194-Q207			D292-S308	V368-L385	D434-Q461		
M5	H82-Y88		E121-D129	E195-Q207	R231-T237		E290-S308	A371-G372	N452-Q461		
M6		Y99-M104	N128-M131	A194-Q207	R231-F240	A254-W260	D303-G309			T505-A512	
M7	H82-P94	Y99-M104	N128-M130	A194-A206	S232-T237	A253-P266	D303-S308	V369-F380	N451-F464	T505-L508	A512-I514

Table 1. Comparison of regions that showed significant structural differences between mutants and regions with higher fluctuations in WT CYP27A1.

Structural alignments between WT and mutant conformers showed no major structural differences for M1-M4 and M7 (RMSD < 1.196 Å), except for M5 and M6 (RMSD = 1.306 and 1.271Å, respectively). In all models, there are common regions that show differences with respect to the native WT conformation. That is, loops between D-E helix (A194-Q207) and G-H helix (L290-V307) (Fig. 5). These regions exhibit high fluctuations even in the WT protein. Prosser et al., reported that the conserved D303 formed a hydrogen bond with residues of the G helix[14]. Only in the native conformation of the WT, a hydrogen bond between D303 and S317 was observed, in all the MPCC this interaction disappears and instead the interaction is observed between Q306 and H312 (3.2 Å) in M3, and with R164 in M5 and M6 (3.1 Å), none of the residues are placed in G helix (Supplementary material Fig. **S11A**) .

In at least 60% of the analyzed CYP27A1 mutants, significant structural differences were observed in regions forming the active site entrance: A helix (H82-P94), loops between β−1a-β−1b sheets (S98-M104), B-B' helix (E121-M131), F-G helix (N250-F267) (Table 1; Fig. 5). Other critical regions included loops between E-F helix (Q230-F240), and J-K helix (V368-L385), which stabilizes the β3 and heme binding region, essential for catalytic activity[35], and the meander region (F438-R459), which stabilize the ERR triad region. The meander region spanning about 20 residues, is proposed as a redox partner interaction site in CYP450 enzymes[36] and shows structural modifications due to mutations. In the native conformation and in the MPCC of the WT, a network of hydrogen bond is observed, E392, R395, R448, and W449 maintaining the union between the meander region and the K helix (Fig. **S11A**). In the MPCC of WT, R395 forms a hydrogen bond with E392 as in the native conformation but the interaction with R448 was not observed. The same behavior is also seen in M1 and M5-M7. However, in M2-3 where mutations of R395C, and R395S are present, the interaction was disrupted and it was was partially offset by interaction with R448 and W449 (Fig. **S11A**), thus the structure was destabilized affecting the meander region and K helix movement. In M1 (G145E), the change of G by E disrupts a hydrogen bond network. In all MPCC, R137 forms a hydrogen bond with E331 which in turn interact with T148 by hydrogen bond. However, in M1, R137 interacts with E145 and the network with E331 is broken, and because of that the fluctuations of loop between B'-C helix are increase (Fig. **S11B**).

Structural modifications in the substrate channel access

Point mutations caused structural changes affecting the protein surface and access channels to the catalytic site. In the initial WT CYP27A1 model, two entrances to the catalytic site were identified (Fig. 6). Channel **I**, is shorter with a wider bottleneck (13.66 × 2.66 Å), Channel **I**, also known as pw2a, is surrounded by loops: F-G, K-β−3a, B-B'; and β-strands: β−1a, β1-1b, β−3b, β−5a, and β−5b. Channel **II**, is larger with a narrower bottleneck (17.09 × 2.090 Å) (Fig. 6A), is formed by E-, F-, I- helices, β−5a and β−5b sheets. Channel **I**is common in the CYP family[37], while channel **II**may correspond to the solvent access channel[14,37].

In contrast, in the MPCC of WT, the lengths of channels **I** and **II** increased, while the bottle neck radius decreased (Figs. 6B and 26.79 × 1.63 and 23.97 × 2.08 Å, respectively). Similar values was observed when the substrate was bound to CYP46A1 (≈ 25 Å), a human cholesterol-24-hydroxylase that shares 34% identity[38].

Comparing mutant channels to the WT-MPCC, channel **I** increased in length in M1 around 1.8 Å (Fig. 6C), while M2-M4 reduced in size by 4–10 Å (Fig. 6D-F), and in M7 by 1.7 Å (Fig. 6I). M5-M6 remained nearly the same size (Fig. 6G-H). Channel **II** reduced in length in M1, M3, M4 and M5 between 1.0 and 4.6 Å (Fig. 6C, E-G), with no changes in M2, and M6 (Fig. 6D **and H**), and increased by 5 Å in M7 (Fig. 6F). Bottleneck radius changes in channel **I** were subtle with all mutants increasing by around 0.20–0.78 Å, especially M1, M2, and M5. In Channel **II**, **the** trend was different; only M1-3 and M6-7 increased in size around 0.14–0.82 Å, while M4 and M5 decreased by 0.3 Å (Fig. 6, **supplementary material (not bol) Table S1**).

Notably, M2 exhibited a third channel (channel **III**), with an intermediate length and smaller bottleneck radius compared to channels **I** and **II** (20.98 and of 1.76 Å, respectively) (Fig. 6D). Channel **III** is formed by B'-, F-, G-, β−5a, and β−5b helices, and is mainly composed of hydrophobic residues. This transient channel has not been previously reported, but its hydrophobic nature suggests it could house cholesterol or other hydrophobic substrate due to its nature and dimensions. Further studies are needed to explore its potential role. Therefore, the MD simulations studies carried out help determine the structural changes in the substrate channel access due to point mutations, which may explain the lack of or the decrease in the enzyme activity.

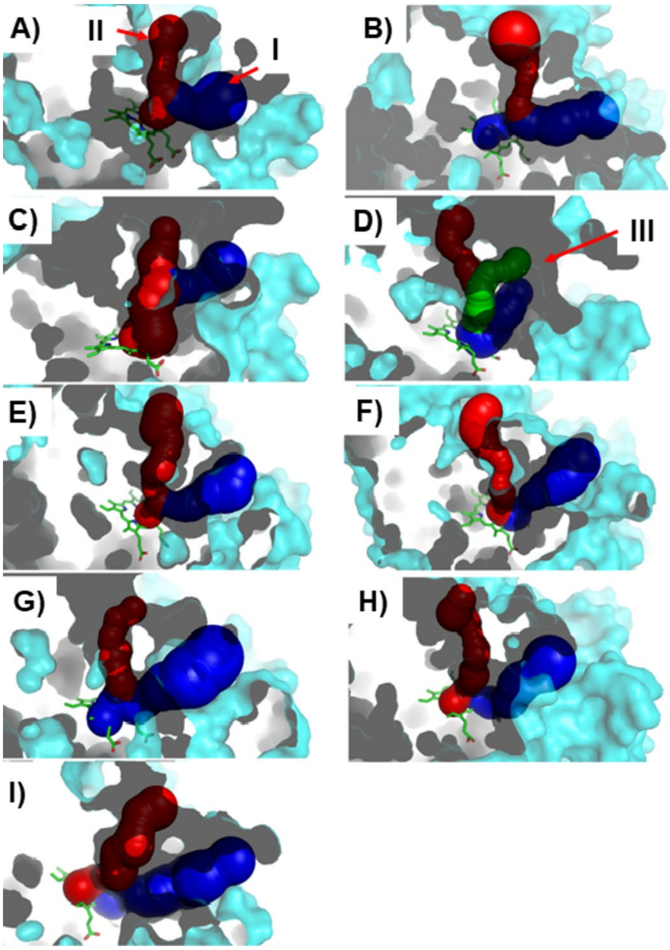

Fig. 6. Structural changes observed in substrate access channels in the WT and mutants identified by Caver 3.0. **(A)** The initial model of WT CYP27A1 shows 2 channels, **I** and **II**. From this point on, the channels were also monitored in the MPCC of **(B)** WT, **(C)** M1, **(D)** M2, **(E)** M3, **(F)** M4, **(G)** M5, **(H)** M6, and **(I)** M7.

Molecular docking suggests modification of the substrate-binding mode

Point mutations produced structural modifications prompting the use of molecular docking to investigate cholesterol binding. In the WT protein, cholesterol reached the catalytic site through hydrophobic channel **I** (Fig. 7A-B). The C26 methyl of cholesterol was inserted into the catalytic site, 3.1 Å from the heme iron. The steroid moiety was accommodated along channel **I**, with the hydroxyl group forming hydrogen bonds with Y253 and T255 resulting in an energetically favorable binding with a free energy of binging of −9.17 kcal/mol (Table 2).

Mutants showed modifications in cholesterol binding compared to WT. In M1 and M2, cholesterol was deeply inserted into Channel **II** (4.2 Å and 3.08 Å away from Fe, respectively), but flipped, orienting the hydroxyl group to the heme iron and the aliphatic chain into the catalytic pocket, with W214 interacting by non-bonding interactions with the aliphatic chain (Fig. 7C-D). Despite favorable binding affinities Table 2), cholesterol would not be hydroxylated, but more specific studies need to be done using the compound I (CpdI), which is the active oxidizing specie that reacts with the substrate[39].

In M3 and M4, cholesterol did not reach the catalytic site (Fig. 7E-F). Interestingly, in M3, it was positioned on the protein surface, 10.5 Å far from the Fe of the heme group. In M4, the aliphatic chain was inserted in a small portion of channel **II**, 13 Å from the heme, with the sterol moiety on the protein surface. In both cases, cholesterol exhibited favorable binding free energy values (Table 2).

Cholesterol in M5-7 also did not adopt a correct binding mode allowing hydroxylation, despite favorable energies (−9.19 to −10.14 kcal/mol, Table 2). The sterol group was closer to the heme group (4.6, 7.0 and 4.6 Å, respectively). In M5, channel **I** was deformed, increasing its length and bottle neck radius accommodating the sterol and protruding the aliphatic chain into the reduced channel **II** (Fig. 7G). In M6, the scaffoldgroup was housed along channel **I** and in a wide subpocket formed by hydrophobic residues (M130, F147, D338-T339, V399-S404, I514, and V515) (Fig. 7H). In M7, the sterol was accommodated similarly to M5, but the aliphatic chain was better accommodated in channel **II** because it increased its dimensions (Fig. 7I). The modifications in binding mode of cholesterol follow a similar trend in M5, M6 and M7 in whose case mutations were introduced at R474 which is a residue that forms a salt bridge with propionate of the Heme group, these modifications

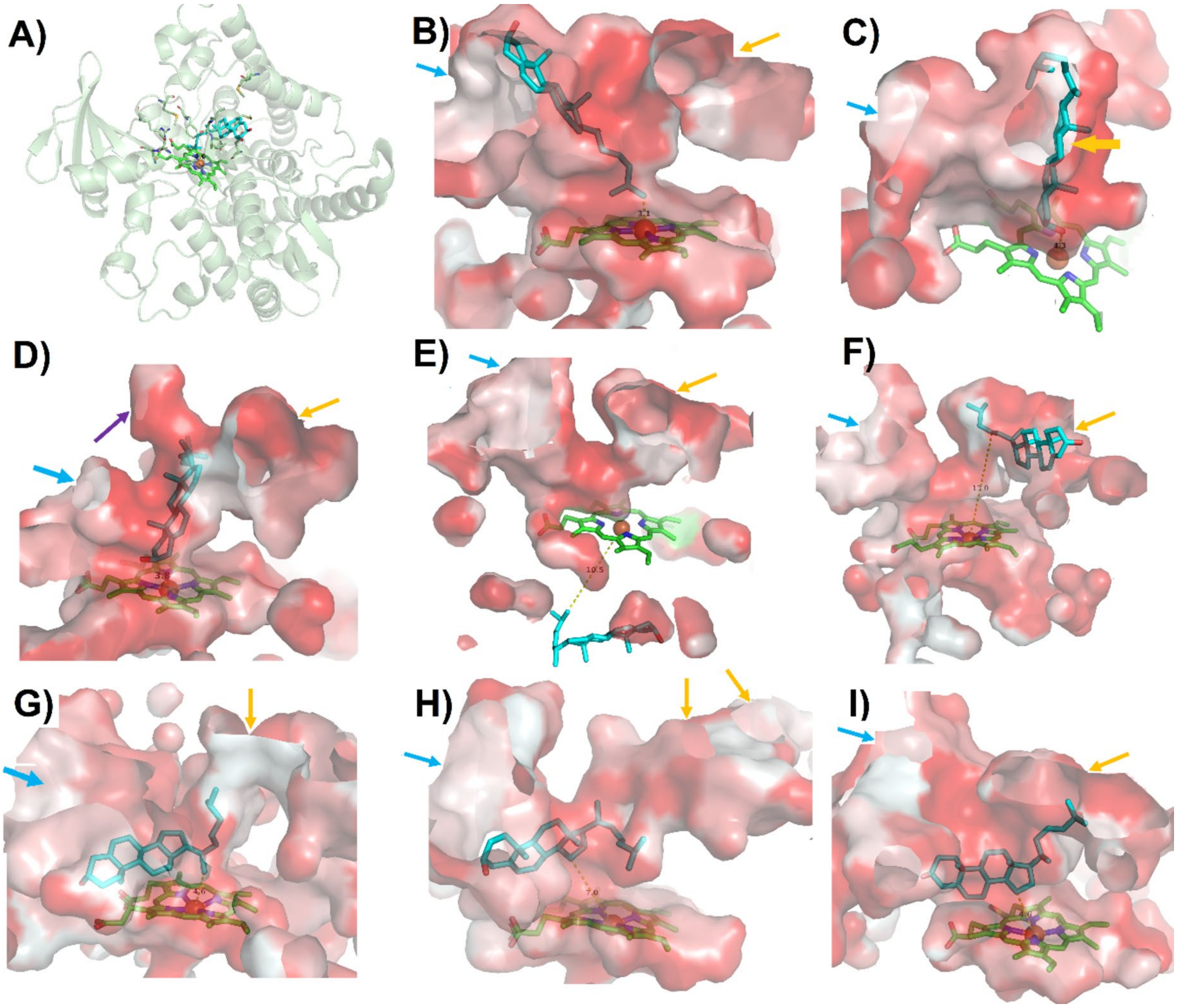

Fig. 7. Binding mode of cholesterol in WT and mutant proteins determined by molecular docking in MPCC. **A**) Ribbon and **B**) surface representation of WT, **C)** M1, **D)** M2, **E)** M3, **F)** M4, **G**) M5, **H**) M6, and **I**) M7 in complex with cholesterol. Heme group is depicted as green sticks and Fe atom as orange ball. Cholesterol is depicted as cyan sticks. Surface representation is colored according to Eisenberg's scale of hydrophobicity[68], where the hydrophobic to non-hydrophobic gradient goes from red to white color. Blue, yellow, and purple arrows point out to channel **I**, **II**, and **III**, respectively.

CYP27A1 model	Free energy of binding (Kcal/mol)	Hydrogen bond	Hydrophobic interactions
WT	−9.17	Y253, T255	D129, M130, F147, L257, L132, W260, Y261, I514, A335, V400, A335, T339, HEME
M1	−8.99	W133	F147, M247, Y253, S251, E331, M334, A335, T339, V400, I514, HEME
M2	−11.59		H136, W133, R137, F147, F279, E331, A335, V400, W275, M334, I514, V515, HEME
M3	−8.10		L386, K3387, L390, A482, E483, L484, E485, M486, Q487, L488
M4	−9.25		F248, Y253, D338, S341, N342, R513, V515, V517, P518, D521, V522
M5	−9.67	S404	W133, F147, R513, Y253, A335, D338, T339, T402, N403, I514, HEME
M6	−9.19	T402	M130, F147, N342, V400, N403, S404, D338, T339, V399, I514, V515
M7	−10.14	T402	R127, W133, M247, F248, W275, M334, A335, D338, T339, V400, N403, S404, V515, HEME

Table 2. Molecular Docking scores and interactions between cholesterol and WT CYP27A1 and the mutants studied.

produce not only an stabilization of Heme group but also a deformation of the Channel I and II leading to modification in the binding pose of Cholesterol inside the channel. In the case of M7 (R479), this mutation also produces a deformation of the channels in similar way to R473, probably because it is located at the thiolate loop an important region connected with C476. It is reported that this region participates in a network of hydrogen bonds that stabilize the heme binding region[14].

Molecular docking allowed us to elucidate the binding mode of the cholesterol in WT CYP27A1 and mutants. These predictions suggest that the complexes are energetically favorable. However, mutants could not properly accommodate cholesterol for 27-hydroxylation because the point mutations affect the structural stability of the protein deforming Channel I, II as well as the heme binding region. These results provide a possible molecular explanation of what has been observed experimentally, where mutations in G145 (M1), R395 (M2-3), R405 (M4), and R474 (M6-7) do not exhibit hydroxylase activity[3,28,40]. The quantum model of the binding site was used to calculate the corresponding van der Waals surface areas of WT, M1, M2, M5 and M7, which are 4078.4, 3774.0, 4136.1, 3877.0 and 4185.8 Å[2], respectively. Although, these values follow the same trend as the corresponding Docking-computed binding energies (*vide infra*), they alone cannot explain the variation in the activity of the WT and the mutants. Thus, a more detailed examination of the protein-ligand interactions is necessary for the comprehension of these differences.

Quantum chemical topology analysis

The protein-ligand (CYP27A1-Cholesterol) interaction energies (IEs) computed at the quantum level are shown in **Table S2 (supplementary material)**. The most favorable one is obtained for the WT model (−39.4 kcal/mol). In contrast, the lowest one was calculated for the M2 model (−2.0kcal/mol), which could be a side effect of the reduction in the cluster size that applied because of the SCF convergence problems. Moreover, it has been observed that counterpoise corrections do not necessarily improve IEs[41], and the basis set superposition error could need other types of emendation[42]. Also, relaxation effects of the free protein and ligand are not counted by this model, since it will require the optimization of the full protein. Therefore, although the estimated IEs point to the WT-cholesterol complex as the most stable one in energetic terms, this is not sufficient to explain the lack of hydroxylase activity of the other mutants, as was stated before. A deeper analysis of the change in the non-covalent interactions present in the protein-ligand complexes is a more convenient course of action for rationalizing these differences. With this in mind, a topological analysis of the electron density, $\rho(\mathbf{r})$, of the complexes was carried out within the Quantum Theory of Atoms in Molecules (QTAIM) and the Non-Covalent Interaction index (NCI) frameworks, which have been shown to be useful for the examination of protein-ligand binding processes[43–46].

In QTAIM, specific interactions are recognized by the presence of bond critical points (BCPs), which correspond to saddle points of the electron density $\rho(\mathbf{r})$ and can be covalent or non-covalent in nature. These are further characterized by the value of the Laplacian of $\rho(\mathbf{r})$ in the BCP, which is negative or positive for covalent or non-covalent bonding, respectively. The strength of each interaction is proportional to the value of $\rho(\mathbf{r})$ at the BCP but can be more properly quantified by using the Espinosa-Mollins-Lecomte relationship[47], which states that that interatomic interaction energies are proportional to $0.5 \cdot V_{BCP}$, where the latter corresponds to the potential energy density evaluated at the BCP. Finally, the interacting atoms are connected by bond paths (gradient lines of $\rho(\mathbf{r})$), BP, which facilitate identifying which pair the BCP corresponds to. The QTAIM analysis of the protein-ligand interactions will be focused on analyzing the interactions of the methyl group of C26, which is directly involved in the cholesterol hydroxylation, as well as the OH group because it has been observed that it is relevant for the complex formation with CYP27A1[15]. **Table S3** (**Supplementary material**) shows the values for the electron density, its Laplacian and the interaction energies computed with the Espinosa-Mollins-Lecomte relationship. The latter will be indicated in parenthesis (in kcal/mol) whenever the specific interactions are discussed. A general observation is that all the examined interactions have positive Laplacian values, confirming the non-covalent nature of these.

Figure 8a and b shows the BCP and BP around the OH and C26 methyl groups of cholesterol in its complex with WT. C26 methyl group form two H···H hydrogen-hydrogen bonds[48], with V400, another one with F147, as well as a H···C and a H···N interaction with the heme group. The corresponding interaction energies are within the interval of 1.2 to 2.0 kcal/mol. Interestingly, the neighbor C27 methyl group (also belonging to the isopropyl group) shows a strong C-H···Fe interaction with the HEME group (5.0 kcal/mol). The H···Fe distance obtained from the partial geometry optimization is 2.2 Å, which is significantly lower than that guessed from the Docking calculation (3.1 Å), which provides additional evidence of the strength of this interaction. This is notorious in the deformation density map (Fig. 8C), where some charge transfer from the iron atom to the interacting hydrogen atom occurs, which is observed as an extension of the negative (blue) region surrounding the Fe. The strong C27-H···Fe contact, as well as the H···H of the C26 methyl group are corroborated by mapping the electron density on the Hirshfeld surface to the cholesterol molecule embedded in the WT binding site (Fig. 8D). The red regions in this surface correspond to zones where electron density is increased because of strong intermolecular interactions. Given that the isopropyl group can rotate easily, it is plausible that that C26 methyl group can also form strong interactions with the Fe atom of heme. We propose that these interactions between isopropyl and heme, which are absent in the other mutant complexes, are responsible for the activation and further oxidation of cholesterol. On the other hand, the OH group of cholesterol act as donor in two strong O-H···O hydrogen bonds with Y253 (5.8 kcal/mol) and T255 (2.7 kcal/mol), and as receptor in two weak C-H···O hydrogen bonds with Y253 (3.2 kcal/mol) and W260 (0.4 kcal/mol). Thus, as mentioned above, these non-covalent forces of the OH group are expected to contribute significantly to the stability and the feasibility of cholesterol oxidation, those with Y253. Notice that interactions involving hydrogen atoms discussed here differ from those described in the previous section. The reason is that in the quantum model the position of H atoms were optimized, meanwhile in the docking simulation they were only inferred.

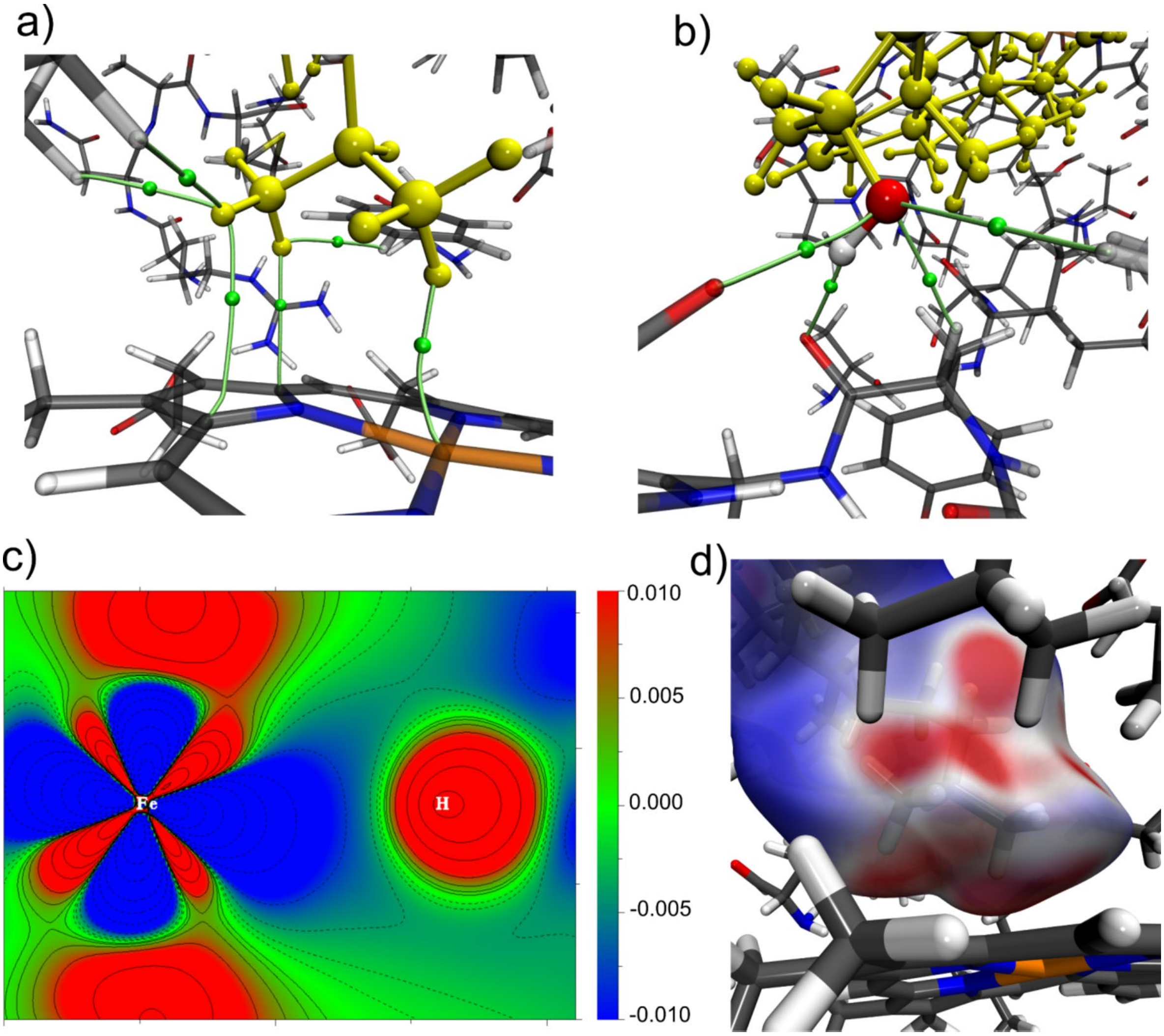

Fig. 8. Bond critical points and bond paths (green dots and lines, respectively) of **(A)** C26 methyl and **(B)** OH groups of cholesterol in its complex with the WT protein. c) Deformation density in the plane formed by the C-H···Fe atoms involved in the cholesterol-heme interaction in the WT complex. **(C)** Electron density mapped on the Hirshfeld surface of cholesterol in the WT complex. Carbon, hydrogen, oxygen, nitrogen and iron atoms are shown in gray, white, red, blue and orange colors, respectively. The aliphatic part of cholesterol is shown in yellow.

In M1 **Fig. S12** (**supplementary material**) C26 methyl group is only involved in one H···H interaction with F248 (3.0 kcal/mol) and another with the OH group of S251 (2.1 kcal/mol), while the OH group forms two very strong interactions with the HEME group; a strong O→Fe coordination bond (5.9 kcal/mol) whose length is 2.63 Å, and a O···N (4.1 kcal/mol) which has a dispersive origin (*vide infra*). Moreover, it acts as an H donor in a very strong O-H···O hydrogen bond (7.0 kcal/mol) with A335, which is the strongest interaction found. Nevertheless, in agreement with our previous hypothesis, despite the formation of these favorable interactions of the hydroxyl group, the C26 methyl group is poorly activated (i.e., its hydrogen atoms form only weak interactions) and thus the hydroxylation reaction is not expected to proceed. A similar result is found for M2, M5 and M7 where, as discussed earlier, the C26 methyl group is far from HEME, and establishes weak interactions only. In M2, it forms several H···H interactions with F279, W275, M334 and E331 ranging from 0.5 to 2.8 kcal/mol, whereas the OH group forms a C-H···O hydrogen bond (4.6 kcal/mol) with V400. In M5, the C26 methyl group forms two H···H interactions with D338 (1.1 kcal/mol) and W133 (0.2 kcal/mol), and two C-H···O hydrogen bonds with Y253 (0.2 kcal/mol) and M334 (2.5 kcal/mol), and the OH group forms a C-H···O hydrogen bond with S404 (2.4 kcal/mol). Finally, in M7 the C26 methyl group forms three H···H interactions with D338 (1.0 kcal/mol), M247 (0.5 kcal/mol) and F248 (0.5 kcal/mol), two C-H···O hydrogen bonds with M334 (2.2 kcal/mol) and D338 (2.2 kcal/mol), and a H···C interaction with M334 (1.1 kcal/mol), while the OH group forms an O-H···O hydrogen bond with T402 (1.0 kcal/mol), an O···N contact with N403 (1.1 kcal/mol), and two C-H···O hydrogen bonds with N403 (0.8 kcal/mol) and T402 (3.5 kcal/mol). In sum, in each mutant the chemical environment surrounding these two important groups is fundamentally different.

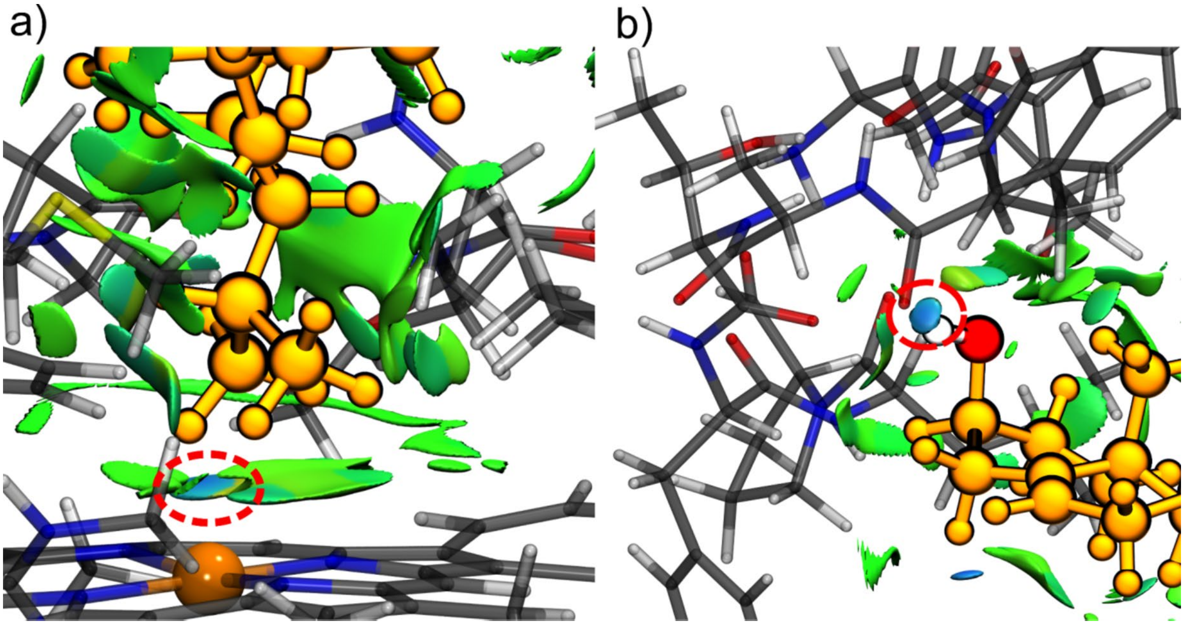

Fig. 9. NCI isosurfaces (0.3) of the cholesterol-WT complex around **(A)** the isopropyl and **(B)** OH groups of the ligand. The strong localized interactions are circled in red. Carbon, hydrogen, oxygen, nitrogen and iron atoms are shown in gray, white, red, blue and orange colors, respectively. The aliphatic part of cholesterol is shown in yellow.

While QTAIM is adequate for describing localized interactions, delocalized ones (such as van der Waals forces) are more properly described by the Non-Covalent Interaction Index (NCI). The NCI approach is based on studying reduced density gradient isosurfaces, which is a quantity that indicates the deviation ρ(**r**) from the uniform electron gas[49]. The shape of these isosurfaces indicate the presence of localized interactions (such as hydrogen bonds), which are manifested as thick disks, or delocalized interactions where more than two atoms are involved (dispersion forces), that are observed as flat extended surfaces. A code color based on the second eigenvalue of the electron density Hessian matrix (λ_2), whose sign indicates charge density accumulation (negative) or depletion (positive), is employed to characterize the strength of the interactions. Blue, green and red are associated to strong, weak and repulsive interactions, correspondingly. Finally, integration of ρ(**r**) in the NCI interaction regions, at different sign(λ_2)·ρ(**r**) intervals can be used to quantify the contribution of the distinct types of interactions to the stability of the protein-ligand complex[50]. Recommended sign(λ_2)·ρ(**r**) intervals (given a.u.) are − 0.1 to −0.02for strong attractive interactions, from − 0.02 to 0.02 for weak interactions, and from 0.02 to 0.1 for repulsive interactions[51].

The NCI isosurfaces of the WT and mutant complexes are shown in Fig. 9 and **Fig. S13** (**Supplementary material**). In the WT complex, it is seen that the H···Fe interaction formed between the isopropyl and HEME groups corresponds to a strong localized interaction (blue region circled in red in Fig. 9a). The rest of the aliphatic part of cholesterol is attached to the protein mainly by dispersive forces, including the H···H interactions which are embedded in green surfaces corresponding to van der Waals regions. The same holds for the interactions of the C26 methyl group of M1, M2, M5 and M7 found in the QTAIM analysis. These are delocalized weak interactions, except for the O-H···H hydrogen-hydrogen bond with S251 in M1, the H···H contact with E331 in M2, and the C-H···O hydrogen bond with M334 in M5, which appear as thick disks of green or blue color, indicating that these are localized interactions of medium or weak strength, respectively. The strong O-H···O hydrogen bond formed between the OH group of cholesterol and Y253 in the WT complex is also observed as an intense-blue thick disk (circled in red in Fig. 9b), while the rest of the interactions described above have a more dispersive nature. In M1 the O→Fe dative bond between cholesterol and the heme group appears as a localized blue region merged with a flat extended van der Waals surface, while the very strong O-H···O hydrogen bond is recognized as a small intense-blue disk. In contrast, in M2, M5 and M7 the interactions of the OH group of cholesterol are mostly of weak and delocalized nature. Finally, from the NCI integrals **Table S3** (**supplementary material**) it is observed that in all complexes dispersive forces predominate since their contribution goes from 63.4% in M1, to 80.0% in M7. In contrast, the repulsion component is the lowest in all complexes (between 5.5 and 11.9%). The complex with the largest contribution from strong localized interactions is M1 (24.9%), which is attributed to the O→Fe coordination bond, although this effect is contrasted by steric repulsion (11.6%).

From the quantum chemical analysis, it is observed that although dispersion forces have a dominant role in the formation of the cholesterol-protein complexes, the binding mode of the cholesterol and the presence of some specific non-covalent interactions are crucial to explain the difference in the activities of WT and the mutants. In particular, it is concluded that the strong interaction between the isopropyl group of cholesterol and the HEME group of WT is essential for the hydroxylation capacity of CYP27A1. Since this interaction was only found in the cholesterol-WT complex, its absence in the other mutants could plausibly explain their lack of activity.

Conclusions

In this research, we explored the structural and dynamic behavior of wild-type (WT) CYP27A1 and seven clinically reported point mutations associated with Cerebrotendinous Xanthomatosis (CTX). Using molecular dynamics (MD) simulations, docking, and quantum calculations, our findings suggests that mutations in CYP27A1 lead to structural changes, which in turn produces affectations in the surfaces of the catalytic site and substrate access channel affecting the binding mode of the natural substrate, cholesterol in CYP27A1. These modifications offer a possible explanation to the reduction in the catalytic activity of the enzyme observed in the progression of CTX. Increased fluctuations were observed in regions around the substrate access channel and the catalytic site (e.g. A helix, loops between E-F, G-H Helix, β sheet-3b-L helix). The increase in the fluctuations in regions around the substrate access channel impact in the structural topology of the channel which become wider or narrower depending on the specific mutant affecting how the substrates enter to the catalytic site. Other critical regions such as loops between E-F helix, and J-K helix which stabilizes the β3 and heme binding region also impact in the catalytic site. According to our MD simulation and docking studies cholesterol was able to reach the catalytic site and adopt the correct binding mode only in WT CYP27A1. Quantum mechanics calculations suggest that the interaction between isopropyl group of the cholesterol and the heme group is essentiall for the hydroxylation capacity of CYP27A1.

Further studies are needed to explore the substrate-CYP27 binding with long simulation times and employing the CPDI, cation radical specie that properly reacts with the substrate.

Methods

Molecular modeling

The primary sequence of CYP27A1 was retrieved from the UNIPROT database (code: Q02318). Furthermore, the tridimensional (3D) model of wild-type CYP27A1 and mutants were separately obtained using the Swissmodel server (https://swissmodel.expasy.org/) by homology modeling[19]. The protein sequence of the mutants studied were aligned with the WT CYP27A1 sequence with Clustal X[52] (Fig. S2 supplementary material).

The crystal structure of rat mitochondrial P450 CYP24A1 was used as a template (PDB: 3K9Y)[20]. The protein sequence of CYP27A1 and the CYP24A1 (UNIPROT code: Q09128) were aligned using the Praline web server to determine the conservation of the amino acids(See alignment Supplementary material Figure **S1**). The calculation of identity and similarity of CYP27A1 and CYP24A1 were done with Bioedit[21,53]with the BLOSUM 62 matrix which is used for comparing protein sequences with identity lesser than 62%[54]. Further, an structural alignment for comparing both structures were done with Chimera 1.16[55]. The quality of the models was verified using the MOLPROBITY server (http://molprobity.biochem.duke.edu/)[56] by validating its dihedral angles using the Ramachandran plot (Supplementary material Fig. **S3-S10**). Further, the generated models were submitted to molecular dynamics (MD) simulations.

Molecular dynamic simulations

The model of WT protein and the mutants (M1-M7) were submitted to MD simulations to refine the structure as well as to study the dynamic behavior of the apo systems. The MD simulations were performed using AMBER 16[57], using the ff14SB force field[58]. The heme parameters of CYP450 were taken from previous research[59]. The systems were prepared using the module Tleap, they were positioned at a rectangular box of 12.0 Å and solvated on explicit TIP3P water[60]. Further, the systems were neutralized by adding the corresponding counterions, specifically Na + around the proteins at physiological pH. The titrable amino acids were verified using Propka[61], and it was not necessary to change the protonation of any residues. Each system was minimized using 2500 steepest descent steps, and 2500 steps of conjugate gradient, restraints were applied to the protein at 100 kcal/mol*$Å^{-2}$ to allow the minimization of the water and ions molecules. Further, the system was submitted to an equilibration protocol as follows: 500 picoseconds (ps) of heating and 500 ps of density equilibration with weak restraints applied to the protein (100 kcal/mol*$Å^{-2}$) and finally, 2 nanoseconds (ns) of equilibration at constant pressure and 310 K with no restraints. The production of the MD simulations was run for 25 ns without position restraints under periodic boundary conditions and with an NPT ensemble. The particle mesh Ewald method was used for describing the electrostatic term[62]; a 10.0 Å cut-off was used, a similar radio was chosen for van der Waals interactions. For controlling the temperature at 310 K, the Langevin Thermostat was used, and for maintaining the pressure at 1 atm, the Berendsen barostat weak-coupling algorithm was employed[63]with coupling constants τT and τp of 1.0 and 0.2 ps, respectively. For constraining the bond length at their equilibrium values the SHAKE algorithm was employed[64], The timestep was set up to 2.0 fs.

Trajectories were analyzed using the cpptraj module of AmberTools 16. The root mean square fluctuation (RMSF), root mean square deviation (RMSD), and radius of gyration (Rg) were calculated. The Clustering analysis was performed using a 3.0 Å of cut-off and a hierarchical agglomerative (bottom-up) approach with AmberTools 16 to obtain the most populated cluster conformation (MPCC) of each mutant and WT protein. Further, the MPCC of each system was evaluated in ProteinTools server (https://proteintools.uni-bayreuth.de/bonds/) for determining a hydrogen bond network[65]and also in FindHBond module of Chimera UCSF[55] for determining possible Hbond.

Structural analysis and visual inspection was performed using PyMOL v0.99[66].

Substrate access channel

To study the structural changes occurring in the substrate access channel of the MPCC of the WT and mutants after the MD simulations, Caver 3.0.3 was used[67]. Caver was used under the following parameters: minimum probe radius of 0.9 Å, a shell depth of 5.0 Å, a shell radius of 7.0 Å, and clustering threshold of 3.5, using the coordinates of Fe as the starting point.

To reveal the hydrophobicity of the substrate access channel, the protein surface was colored using the script color_H.py from the University of Osaka. The hydrophobicity scale is based on Eisenberg's scale[68], where the hydrophobic to non-hydrophobic gradient goes from red to white color.

Molecular docking

The most populated cluster conformation obtained from the MD simulations of each of the studied models was employed in the molecular docking studies using cholesterol as the substrate.

The receptors and ligands were prepared using AutoDock Tools 1.5[69]. For ligands Gasteiger charges were assigned and only polar hydrogens were retained. While, for the proteins Kolman charges were calculated and only polar hydrogens were assigned. We used focused docking as part of our approach so that, grid box was on the heme group with a size of 70 Å x 70 Å x 70 Å with a grid space of 0.375 Å[3]. The scoring sampling used was the Lamarckian Genetic Algorithm with a randomized population of 100 individuals and a maximum number of energy evaluations of 1×10^7; in total 100 runs were performed for each system. Finally, molecular docking was run in Autodock4 [69]. Of each docking only a representative conformer of the most populated cluster with the lowest energy binding was retained for further quantum studies.

Quantum calculations

The Docking-obtained geometries were taken as input for generating the quantum models. All amino acids (as well as the complete HEME group) lying in a 6 Å radius around the cholesterol molecule were included in the model of all the analyzed systems. This radius was chosen because no appreciable difference was observed in the integration of the electron density in the NCI isosurfaces if larger radii were considered. The terminal NH and CO groups were completed with hydrogen atoms. A constrained geometry optimization, using the GFN2-xTB semi empirical method was performed with the *xtb* program (https://github.com/grimme-lab/xtb/releases) (6.6.1 version)[70]. The full cholesterol molecule, as well as all the H and Fe atoms were free to move, while the rest of the system was constrained by applying a harmonic potential with a 0.1 a.u. force constant. The Non-Covalent Interaction index (NCI)[71]was used to analyze the protein-ligand intermolecular forces within the pro molecular electron density approach. This analysis was carried out with the NCIPLOT program[72] (version 4.2), taking the GFN2-xTB-optimized geometry as input coordinates. Also, to characterize the nature and strength of intermolecular contacts, the electron density, $\rho(\mathbf{r})$, of the protein-ligand complex was studied by means of the Quantum Theory of Atoms in Molecules[73] (QTAIM). The $\rho(\mathbf{r})$ was determined by single point calculations on the GFN2-xTB-optimized geometries, at the PBE0/6-31G(d, p) level of theory with the Orca software[74](version 5.0.3). This level of theory was selected because it brings an optimal cost/benefit relation, in relation to the description of the electron density, for which this hybrid functional provides accurate results[75]. For the M2 and M7 mutants, the amino acids that were not directly interacting with the cholesterol molecule were dropped out of the DFT calculations because of problems related to the Self Consistent Field convergence. Additionally, the deformation density (defined as the difference between the actual $\rho(\mathbf{r})$ and one constructed as a sum of spherical non-interaction atoms) was analyzed to provide further evidence of intermolecular interactions. This method can assess charge flow caused by the presence of all types of chemical interactions[76]. Besides, the $\rho(\mathbf{r})$ mapped on the Hirshfeld surface of the cholesterol molecule embedded in the binding sites of the proteins was also studied. This function has become important for visualizing regions of close contacts in supramolecular chemistry[77]. The QTAIM analysis, deformation density and Hirshfeld surfaces were performed with Multiwfn software[78](version 3.8). Finally, the interaction energies were computed as the difference between the protein-ligand complex and the free protein and ligand energies at the same geometry as in the complex, with the same level of theory. Dispersion and basis set superposition error corrections were added to all energies using the DFT-D3 method[79]with the Becke-Johnson damping function[80], and the geometrical Counterpoise scheme[81], respectively.

Data availability

All relevant data supporting the conclusion of the article is included in the manuscript.

Received: 23 September 2024; Accepted: 11 March 2025

References

1. Sundaram, S. S., Bove, K. E., Lovell, M. A. & Sokol, R. J. Mechanisms of disease: inborn errors of bile acid synthesis. *Nat. Clin. Pract. Gastroenterol. Hepatol.* **5**, 456–468. https://doi.org/10.1038/ncpgasthep1179 (2008).
2. Griffiths, W. J. et al. Additional pathways of sterol metabolism: evidence from analysis of Cyp27a1-/- mouse brain and plasma. *Biochim. Biophys. Acta Mol. Cell. Biol. Lipids*. **1864**, 191–211. https://doi.org/10.1016/j.bbalip.2018.11.006 (2019).
3. Sawada, N., Sakaki, T., Kitanaka, S., Kato, S. & Inouye, K. Structure-function analysis of CYP27B1 and CYP27A1. Studies on mutants from patients with vitamin D-dependent rickets type I (VDDR-I) and cerebrotendinous xanthomatosis (CTX). *Eur. J. Biochem.* **268**, 6607–6615. https://doi.org/10.1046/j.0014-2956.2001.02615.x (2001).
4. Mast, N. et al. Cytochrome P450 27A1 deficiency and regional differences in brain sterol metabolism cause Preferential Cholestanol accumulation in the cerebellum. *J. Biol. Chem.* **292**, 4913–4924. https://doi.org/10.1074/jbc.M116.774760 (2017).
5. Griffiths, W. J. & Wang, Y. Oxysterols as lipid mediators: their biosynthetic genes, enzymes and metabolites. *Prostaglandins Other Lipid Mediat.* **147**, 106381. https://doi.org/10.1016/j.prostaglandins.2019.106381 (2020).

6. Wong, J. C., Walsh, K., Hayden, D. & Eichler, F. S. Natural history of neurological abnormalities in cerebrotendinous xanthomatosis. *J. Inherit. Metab. Dis.* **41**, 647–656. https://doi.org/10.1007/s10545-018-0152-9 (2018).
7. Stelten, B. M. L., van de Warrenburg, B. P. C., Wevers, R. A. & Verrips, A. Movement disorders in cerebrotendinous xanthomatosis. *Parkinsonism Relat. Disord.* **58**, 12–16. https://doi.org/10.1016/j.parkreldis.2018.07.006 (2019).
8. Salen, G. & Steiner, R. D. Epidemiology, diagnosis, and treatment of cerebrotendinous xanthomatosis (CTX). *J. Inherit. Metab. Dis.* **40**, 771–781. https://doi.org/10.1007/s10545-017-0093-8 (2017).
9. Lee, M. H. et al. Fine-mapping, mutation analyses, and structural mapping of cerebrotendinous xanthomatosis in U.S. Pedigrees. *J. Lipid Res.* **42**, 159–169. https://doi.org/10.1016/s0022-2275(20)31675-8 (2001).
10. Casey, A. & Dolan, L. Genes encoding cytochrome P450 monooxygenases and glutathione S-transferases associated with herbicide resistance evolved before the origin of land plants. *PLoS One.* **18**, e0273594. https://doi.org/10.1371/journal.pone.0273594 (2023).
11. Jones, G., Prosser, D. E. & Kaufmann, M. Cytochrome P450-mediated metabolism of vitamin D. *J. Lipid Res.* **55**, 13–31. https://doi.org/10.1194/jlr.R031534 (2014).
12. Lee, C. W. et al. Clinical and molecular genetic features of cerebrotendinous xanthomatosis in Taiwan: Report of a novel CYP27A1 mutation and literature review. *J Clin Lipidol* 13, 954–959 e951, (2019). https://doi.org/10.1016/j.jacl.2019.10.001
13. Berginer, V. M., Salen, G. & Patel, S. B. in *Rosenberg's Mol. Genetic Basis Neurol. Psychiatric Disease* 589–598 (2015).
14. Prosser, D. E., Guo, Y., Jia, Z. & Jones, G. Structural motif-based homology modeling of CYP27A1 and site-directed mutational analyses affecting vitamin D hydroxylation. *Biophys. J.* **90**, 3389–3409. https://doi.org/10.1529/biophysj.105.069369 (2006).
15. Mast, N. et al. Distinct binding of cholesterol and 5beta-cholestane-3alpha,7alpha,12alpha-triol to cytochrome P450 27A1: evidence from modeling and site-directed mutagenesis studies. *Biochemistry* **45**, 4396–4404. https://doi.org/10.1021/bi052654w (2006).
16. Sunkar, S. & Neeharika, D. CYP2R1 and CYP27A1 genes: an in Silico approach to identify the deleterious mutations, impact on structure and their differential expression in disease conditions. *Genomics* **112**, 3677–3686. https://doi.org/10.1016/j.ygeno.2020.04.017 (2020).
17. Norlin, M., von Bahr, S., Bjorkhem, I. & Wikvall, K. On the substrate specificity of human CYP27A1: implications for bile acid and Cholestanol formation. *J. Lipid Res.* **44**, 1515–1522. https://doi.org/10.1194/jlr.M300047-JLR200 (2003).
18. Shukla, S. & Tekwani, B. L. Histone deacetylases inhibitors in neurodegenerative diseases, neuroprotection and neuronal differentiation. *Front. Pharmacol.* **11**, 537. https://doi.org/10.3389/fphar.2020.00537 (2020).
19. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303. https://doi.org/10.1093/nar/gky427 (2018).
20. Annalora, A. J. et al. Crystal structure of CYP24A1, a mitochondrial cytochrome P450 involved in vitamin D metabolism. *J. Mol. Biol.* **396**, 441–451. https://doi.org/10.1016/j.jmb.2009.11.057 (2010).
21. A., H. T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98. https://doi.org/10.14601/Phytopathol_Mediterr-14998u1.29 (1999).
22. Xiang, Z. Advances in homology protein structure modeling. *Curr. Protein Pept. Sci.* **7**, 217–227. https://doi.org/10.2174/138920306777452312 (2006).
23. Riayn, J. & A., R. in *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing* 455–458 (2009).
24. Hasemann, C. A., Kurumbail, R. G., Boddupalli, S. S., Peterson, J. A. & Deisenhofer, J. Structure and function of cytochromes P450: a comparative analysis of three crystal structures. *Structure* **3**, 41–62. https://doi.org/10.1016/s0969-2126(01)00134-4 (1995).
25. Murtazina, D. et al. Membrane-protein interactions contribute to efficient 27-hydroxylation of cholesterol by mitochondrial cytochrome P450 27A1. *J. Biol. Chem.* **277**, 37582–37589. https://doi.org/10.1074/jbc.M204909200 (2002).
26. Pikuleva, I. A., Puchkaev, A. & Bjorkhem, I. Putative helix F contributes to regioselectivity of hydroxylation in mitochondrial cytochrome P450 27A1. *Biochemistry* **40**, 7621–7629. https://doi.org/10.1021/bi010193i (2001).
27. Coleman, T. et al. Structural insights into the role of the acid-alcohol pair of residues required for dioxygen activation in cytochrome P450 enzymes. *J. Biol. Inorg. Chem.* **25**, 583–596. https://doi.org/10.1007/s00775-020-01781-4 (2020).
28. Cali, J. J., Hsieh, C. L., Francke, U. & Russell, D. W. Mutations in the bile acid biosynthetic enzyme sterol 27-hydroxylase underlie cerebrotendinous xanthomatosis. *J. Biol. Chem.* **266**, 7779–7783. https://doi.org/10.1016/s0021-9258(20)89518-0 (1991).
29. Smalley, S. V. et al. Novel splice-affecting variants in CYP27A1 gene in two Chilean patients with cerebrotendinous xanthomatosis. *Genet. Mol. Biol.* **38**, 30–36. https://doi.org/10.1590/S1415-475738120140087 (2015).
30. Blauwendraat, C. et al. The wide genetic landscape of clinical frontotemporal dementia: systematic combined sequencing of 121 consecutive subjects. *Genet. Med.* **20**, 240–249. https://doi.org/10.1038/gim.2017.102 (2018).
31. Abe, R. et al. Spinal form cerebrotendinous xanthomatosis patient with long spinal cord lesion. *J. Spinal Cord Med.* **39**, 726–729. https://doi.org/10.1179/1079026815Z.000000000409 (2016).
32. Gallus, G. N., Dotti, M. T. & Federico, A. Clinical and molecular diagnosis of cerebrotendinous xanthomatosis with a review of the mutations in the CYP27A1 gene. *Neurol. Sci.* **27**, 143–149. https://doi.org/10.1007/s10072-006-0618-7 (2006).
33. Sasamura, A. et al. Late-onset cerebrotendinous xanthomatosis with a novel mutation in the CYP27A1 gene. *Intern. Med.* **57**, 1611–1616. https://doi.org/10.2169/internalmedicine.0120-17 (2018).
34. Yoshinaga, T. et al. Clinical and radiological findings of a cerebrotendinous xanthomatosis patient with a novel p.A335V mutation in the CYP27A1 gene. *Intern. Med.* **53**, 2725–2729. https://doi.org/10.2169/internalmedicine.53.2996 (2014).
35. Gupta, R. P., Patrick, K. & Bell, N. H. Mutational analysis of CYP27A1: assessment of 27-hydroxylation of cholesterol and 25-hydroxylation of vitamin D. *Metabolism* **56**, 1248–1255. https://doi.org/10.1016/j.metabol.2007.04.023 (2007).
36. Pikuleva, I. A., Cao, C. & Waterman, M. R. An additional electrostatic interaction between adrenodoxin and P450c27 (CYP27A1) results in tighter binding than between adrenodoxin and p450scc (CYP11A1). *J. Biol. Chem.* **274**, 2045–2052. https://doi.org/10.1074/jbc.274.4.2045 (1999).
37. Wade, R. C., Winn, P. J., Schlichting, I. & Sudarko A survey of active site access channels in cytochromes P450. *J. Inorg. Biochem.* **98**, 1175–1182. https://doi.org/10.1016/j.jinorgbio.2004.02.007 (2004).
38. Mast, N. et al. Crystal structures of substrate-bound and substrate-free cytochrome P450 46A1, the principal cholesterol hydroxylase in the brain. *Proc. Natl. Acad. Sci. U S A.* **105**, 9546–9551. https://doi.org/10.1073/pnas.0803717105 (2008).
39. Qureshi, M., Mokkawes, T., Cao, Y. & de Visser, S. P. Mechanism of the oxidative Ring-Closure reaction during gliotoxin biosynthesis by cytochrome P450 GliF. *Int. J. Mol. Sci.* **25** https://doi.org/10.3390/ijms25168567 (2024).
40. Chen, W., Kubota, S., Ujike, H., Ishihara, T. & Seyama, Y. A novel Arg362Ser mutation in the sterol 27-hydroxylase gene (CYP27): its effects on pre-mRNA splicing and enzyme activity. *Biochemistry* **37**, 15050–15056. https://doi.org/10.1021/bi9807660 (1998).
41. Alvarez-Idaboy, J. R. & Galano, A. Counterpoise corrected interaction energies are not systematically better than uncorrected ones: comparison with CCSD(T) CBS extrapolated values. *Theor. Chem. Acc.* **126**, 75–85. https://doi.org/10.1007/s00214-009-0676-z (2009).
42. Jensen, F. Basis set superposition errors are partly basis set imbalances. *J. Chem. Theory Comput.* **20**, 767–774. https://doi.org/10.1021/acs.jctc.3c01156 (2024).
43. Frontera, A. & Bauza, A. Biological halogen bonds in protein-ligand complexes: a combined QTAIM and NCIPlot study in four representative cases. *Org. Biomol. Chem.* **19**, 6858–6864. https://doi.org/10.1039/d1ob01212f (2021).
44. Isert, C., Atz, K., Riniker, S. & Schneider, G. Exploring protein-ligand binding affinity prediction with electron density-based geometric deep learning. *RSC Adv.* **14**, 4492–4502. https://doi.org/10.1039/d3ra08650j (2024).

45. Kubaib, A. et al. Insights into the binding mechanism of 2,5-substituted 4-pyrone derivatives as therapeutic agents for fused dimeric interactions: A computational study using QTAIM, dynamics and Docking simulations of protein–ligand complexes. *Int. J. Quantum Chem.* **124** https://doi.org/10.1002/qua.27330 (2024).
46. Wieduwilt, E. K. et al. Extracting quantitative information at quantum mechanical level from noncovalent interaction index analyses. *J. Chem. Theory Comput.* **19**, 1063–1079. https://doi.org/10.1021/acs.jctc.2c01092 (2023).
47. Espinosa, E., Molins, E. & Lecomte, C. Hydrogen bond strengths revealed by topological analyses of experimentally observed electron densities. *Chem. Phys. Lett.* **285**, 170–173. https://doi.org/10.1016/s0009-2614(98)00036-0 (1998).
48. Matta, C. F., Hernandez-Trujillo, J., Tang, T. H. & Bader, R. F. Hydrogen-hydrogen bonding: a stabilizing interaction in molecules and crystals. *Chemistry* **9**, 1940–1951. https://doi.org/10.1002/chem.200204626 (2003).
49. Boto, R. A., Contreras-García, J., Tierny, J. & Piquemal, J. P. Interpretation of the reduced density gradient. *Mol. Phys.* **114**, 1406–1414. https://doi.org/10.1080/00268976.2015.1123777 (2015).
50. Laplaza, R. et al. NCIPLOT and the analysis of noncovalent interactions using the reduced density gradient. *WIREs Comput. Mol. Sci.* **11** https://doi.org/10.1002/wcms.1497 (2020).
51. Peccati, F. NCIPLOT4 guide for biomolecules: an analysis tool for noncovalent interactions. *J. Chem. Inf. Model.* **60**, 6–10. https://doi.org/10.1021/acs.jcim.9b00950 (2020).
52. Larkin, M. A. et al. Clustal W and clustal X version 2.0. *Bioinformatics* **23**, 2947–2948. https://doi.org/10.1093/bioinformatics/btm404 (2007).
53. Inmunomedicine group, U. C. & d., M. *SIAS: Sequence identities and similarities*, (2008). http://imed.med.ucm.es/Tools/sias.html
54. Pearson, W. R. Selecting the Right Similarity-Scoring Matrix. *Curr Protoc Bioinformatics* 43, 3 5 1–3 5 9, (2013). https://doi.org/10.1002/0471250953.bi0305s43
55. Pettersen, E. F. et al. UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. https://doi.org/10.1002/jcc.20084 (2004).
56. Williams, C. J. et al. More and better reference data for improved all-atom structure validation. *Protein Sci. 27*. **MolProbity**, 293–315. https://doi.org/10.1002/pro.3330 (2018).
57. Case, D. A. et al. The amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688. https://doi.org/10.1002/jcc.20290 (2005).
58. Duan, W. et al. Design, synthesis, and antitumor evaluation of novel histone deacetylase inhibitors equipped with a Phenylsulfonylfuroxan module as a nitric oxide donor. *J. Med. Chem.* **58**, 4325–4338. https://doi.org/10.1021/acs.jmedchem.5b00317 (2015).
59. Shahrokh, K., Orendt, A., Yost, G. S. & Cheatham, T. E. 3 Quantum mechanically derived AMBER-compatible Heme parameters for various States of the cytochrome P450 catalytic cycle. *J. Comput. Chem.* **33**, 119–133. https://doi.org/10.1002/jcc.21922 (2012).
60. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935. https://doi.org/10.1063/1.445869 (1983).
61. Li, H., Robertson, A. D. & Jensen, J. H. Very fast empirical prediction and rationalization of protein pKa values. *Proteins* **61**, 704–721. https://doi.org/10.1002/prot.20660 (2005).
62. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald-an N.Log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
63. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684. https://doi.org/10.1063/1.448118 (1984).
64. van Gunsteren, W. F. & Berendsen, H. J. C. Algorithms for macromolecular dynamics and constraint dynamics. *Mol. Phys.* **34**, 1311–1327. https://doi.org/10.1080/00268977700102571 (1977).
65. Ferruz, N., Schmidt, S. & Hocker, B. ProteinTools: a toolkit to analyze protein structures. *Nucleic Acids Res.* **49**, W559–W566. https://doi.org/10.1093/nar/gkab375 (2021).
66. DeLano. (ed LLC. Schrödinger). (2002).
67. Pavelka, A. et al. Algorithms for analyzing dynamics of tunnels in macromolecules. *IEEE/ACM Trans. Comput. Biol. Bioinform 13*. **CAVER**, 505–517. https://doi.org/10.1109/TCBB.2015.2459680 (2016).
68. Eisenberg, D., Schwarz, E., Komaromy, M. & Wall, R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* **179**, 125–142. https://doi.org/10.1016/0022-2836(84)90309-7 (1984).
69. Morris, G. M. et al. AutoDock4 and AutoDockTools4: automated Docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791. https://doi.org/10.1002/jcc.21256 (2009).
70. Bannwarth, C., Ehlert, S. & Grimme, S. GFN2-xTB-An accurate and broadly parametrized Self-Consistent Tight-Binding quantum chemical method with multipole electrostatics and Density-Dependent dispersion contributions. *J. Chem. Theory Comput.* **15**, 1652–1671. https://doi.org/10.1021/acs.jctc.8b01176 (2019).
71. Johnson, E. R. et al. Revealing noncovalent interactions. *J. Am. Chem. Soc.* **132**, 6498–6506. https://doi.org/10.1021/ja100936w (2010).
72. Boto, R. A. et al. Fast, robust, and quantitative analysis of noncovalent interactions. *J. Chem. Theory Comput. 16*. **NCIPLOT4**, 4150–4158. https://doi.org/10.1021/acs.jctc.0c00063 (2020).
73. Bader, R. F. W. & Nguyen-Dang, T. T. *Adv. Quantum Chem.* 63–124 (1981).
74. Neese, F., Wennmohs, F., Becker, U. & Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **152**, 224108. https://doi.org/10.1063/5.0004608 (2020).
75. Medvedev, M. G., Bushmarinov, I. S., Sun, J., Perdew, J. P. & Lyssenko, K. A. Density functional theory is straying from the path toward the exact functional. *Science* **355**, 49–52. https://doi.org/10.1126/science.aah5975 (2017).
76. Nieuwland, C., Vermeeren, P. & Bickelhaupt, F. M. Fonseca Guerra, C. Understanding chemistry with the symmetry-decomposed Voronoi deformation density charge analysis. *J. Comput. Chem.* **44**, 2108–2119. https://doi.org/10.1002/jcc.27184 (2023).
77. Spackman, M. A. & Jayatilaka, D. Hirshfeld surface analysis. *CrystEngComm* **11**, 19–32. https://doi.org/10.1039/b818330a (2009).
78. Lu, T. & Chen, F. Multiwfn: a multifunctional wavefunction analyzer. *J. Comput. Chem.* **33**, 580–592. https://doi.org/10.1002/jcc.22885 (2012).
79. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate Ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104. https://doi.org/10.1063/1.3382344 (2010).
80. Grimme, S., Ehrlich, S. & Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **32**, 1456–1465. https://doi.org/10.1002/jcc.21759 (2011).
81. Kruse, H. & Grimme, S. A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *J. Chem. Phys.* **136**, 154101. https://doi.org/10.1063/1.3700154 (2012).
82. Laskowski, R. A. et al. Structural summaries of PDB entries. *Protein Sci.* **27**, 129–134. https://doi.org/10.1002/pro.3289 (2018).

Acknowledgements

The authors thank CONACYT-SIP-IPN. BLR acknowledges DGTIC-UNAM (project LANCAD-UNAM-DGTIC-426) for the supercomputer resources.

Author contributions

Y.S.L.: writing, MD simulation, writing, reviewing, data curation; H. L. M.F.: conceptualization, data curation, writing; B.C. L.R.: quantum calculations, writing, reviewing; A. C.M.: conceptualization review, writing; J. C. B.: Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization.

Declarations

Competing interests

The authors declare no competing interests.

Declaration of competing interest

The authors declare there are no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-93966-7.

Correspondence and requests for materials should be addressed to Y.S.-L. or J.C.-B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025