nature portfolio

Peer Review File

Unveiling Cytochrome P450 Enzymes that Catalyze Steroid Side-Chain Cleavage in Bacteria

Corresponding Author: Professor Xudong Qu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The study by Tian et al. present an impressive effort in identifying biotechnologically bacterial cytochrome P450 enzymes that catalyse steroid side-chain cleavage by employing a plethora of biochemical, bioinformatic & biophysical methods.

However, the manuscript lacks some important information, and has presentational issues and draws conclusions from incomplete analyses.

Upon revision, the points given below should be addressed.

Summary:

The study aims at identifying bacterial C14-alpha hydroxylase Cytochrome P450 (CYP) enzymes for biotechnological applications but unexpectedly revealed a bacterial enzyme catalysing steroid side-chain cleavage. These bacterial CYPs bear the advantage of solubility when compared to their membrane-bound mammalian counterparts.

To achieve this, the authors applied an impressive, holistic approach encompassing identification of candidate enzymes by bioinformatic- & biochemical screening, investigation of substrate-specificity & enzymatic mechanism by structural biology, molecular dynamics simulations and biochemical assays & enhancement of enzymatic activity through mutagenesis experiments.

The remarkably wide scope of methods & characterization approaches presented in this manuscript necessitates a wide range of areas of expertise for an appropriate review. This review will, thus, mostly focus on the computational- (docking & MD simulation) and structural sections of the manuscript.

Initially, candidate enzymes were identified from bioinformatic screening and were subjected to experimental screening on different Steroid substrates. Only one catalytically active enzyme: P450-A8 was found to transform cholestenone to Progesterone while also working on a wider substrate range (different from eukaryotic equivalents). This sequence was used to identify 8 further homologous bacterial enzymes for further testing which revealed CYP204A5 to have the highest activity.

To solve the enzymatic mechanism, mono- & dihydroxylated intermediates were identified through comparative LC-MS & chemical synthesis indicating a 2-step hydroxylation mechanism.

This was further supported by structural studies on CYP204A3 (A5 was not crystallizable) & associated docking and molecular dynamics simulations. From the latter and site-saturation mutagenesis experiments, residues facilitating substrate positioning & enzymatic activity were identified yielding novel mutants with increased activity. Further MD studies on 3 representative mutants demonstrated that these mutations place the C20 & C22 atoms closer to the catalytically active Compound 1 species.

Major concerns:

-It is unclear if CYP204A3, whose structure was solved crystallographically, differs from the most efficient candidate identified in this study, CYP204A5, in substrate specificity. The authors only remark on its lower activity, but a thorough comparison of enzyme specificity would be necessary to assess how good of a template this structure is for CYP204A5. Likewise, some measures like sequence identity and sequence alignment would should be provided to assist assessment of the authors' choice to crystallize this enzyme.

-From their comparison of the crystallographic structure of CYP204A3 to mammalian CYP11A1 the authors conclude that the former is less specific due to a looser fit of the docked substrate. This comparison seems very challenging as they compare between a crystallographic holo- and an apo-structure, thus, potentially missing out on induced-fit effects taking place upon substrate binding. It might be necessary to compare to the converged MD-structure of the docked protein-substrate complex, compare quantities like ligand RMSDs or even try to co-crystallize a steroid bound complex to substantiate this point.

-The manuscript lacks any thorough validation of the docking poses of cholestenone to CYP204A3 : what were the docking scores? How do the poses e.g. compare to alignments with CYP11A1:Cholesterole?

Likewise, it would be good to give a proper interaction plot, not only the Surface-fit presented in Fig S8 and to provide the docked structure as supplementary material.

-Likewise, any critical validation of the performed MD simulation is missing. Without any measure of convergence besides the C20, C22-Fe distances, it is impossible to assess if the simulation-lengths & choice to only simulate one replica per system are sufficient. This is very important as some of these distance-plots e.g. in Fig. 6 B & D as well as S11 B & S12 B display increases/instabilities just before the simulations were terminated indicating some structural rearrangements.

Such measures could include e.g. protein & ligand RMSDs over all trajectories.

-In the sentence 'Closer examination of the substrate-binding pocket further identified several key amino acid residues […]' (p. 9) it is unclear how this examination was performed. Was this based on quantitative measurements such as contact analyses or on visual inspection? This information is necessary to assess the validity of all subsequent experiments.

-The 'Methods' & 'Supplementary methods' sections lack essential information on the simulation-based assays:

-Which parameters were used in SwissDock ?

-Which implicit solvent model was used in DFT?

-Which spin-state was chosen to represent CPD1? How high was the associated spin-contamination? High spin-contamination can over-delocalize charges over the molecule, thus, affecting RESP calculations.

-How were the RESP charges computed in Multiwfn? Which fitting algorithm etc. was utilized? How were the charges made compatible with AMBER?

—> why did the authors choose to use Multiwfn for RESP charge computations instead of the native AMBER program Antechamber which can also compute RESP charges from ESP calculations with Gaussian. Likewise, why was the 'sobtob' tool utilized instead of Antechamber

--> How was the system protonated and at what pH?

-> References for the AMBER14SB-PARMBSC1 and GAFF force fields and the TIP3P water model should be given.

-How many counter-ions were added to the simulation systems, i.e. at which concentration?

-No information on electrostatics, treatment of periodic boundary condition etc. are given for the energy minimizations

-The 'Methods' section of the manuscript essentially only describes experimental methods. Despite majorly contributing to the manuscript, almost all bioinformatics- and simulation-related methods have to be searched for in the Supplementary materials. Especially in the light of our concerns about critical information lacking there, it is unclear to us why the methodology was fragmented like this.

Minor concerns:

-Language: some incomprehensible sentences, some repetitions/tautologies (e.g. '[…] plant CYP87A family P450scc enzymes […] have recently been identified in several plants'); many titles of Supplementary materials erroneous

-The authors' initial screening approach seems to undervalue the importance of local vs. global similarity. The way we understand their initial bioinformatic approach, it was based essentially on just selecting one template enzyme and a global alignment to identify initial hits. Since CYPs are very well known & relatively conserved, it might have been more successful if the authors would have instead focused on local alignments in the known substrate-coordinating regions of steroid-metabolizing CYPs. This point should be elaborated on in the introduction and/or discussion

-The manuscript's clarity would greatly profit from using proper abbreviations instead of numbers to refer to different compounds e.g. PROG/PREG/…, particularly as this gets very confusing when reaction steps are discussed.

-The text in Fig. 2 is too small and not readable

-Fig. 4 E: the presented mechanism does not seem to make sense: where does the third oxygen in 'Cycle 3' come from? Is it derived from a water molecule as in Cycle 3, CPD1 still carries an oxygen?

-Actual means & standard deviations for the distances measured over the MD simulation should be reported instead of discussing 'approximate values' in paragraph 2 on p. 9

-The title of table S2 is entirely unclear to us ('Information of redox used in this study')

-In the Supplementary methods, the authors refer to AF3 models as homology models which is wrong.

-Why did the authors choose the AMBER14SB-PARMBSC1 force field?

Reviewer #2

(Remarks to the Author)
The manuscript by Tian et al. describes the discovery of bacterial P450scc enzymes in steroid-degrading microorganisms. The authors reveal that these bacterial enzymes proceed via a different mechanism compared to the mechanism of the comparable enzymes found in mammals and plants. The authors also perform engineering work based on one obtained bacterial P450scc enzyme structure. Overall this is an excellent paper, providing new insights into bacterial metabolism and also highighting a new biocatalytic tool for production of progesterone-derived compounds.

Below are minor commenst:

1) The stereochemistry of the compounds in Fig. 1 is neither complete (1, 2 ,3, and 6), nor correct (Fig. 1C, intermediate 6).
2) The intermediates in Fig. 1. are numbered wrongly.
3) Please make Fig. 3B more reader friendly since some annotations are upside-down.
4) Please explain why the peaks for 5 in Fig. 4D are not consistent by retention time.
5) Figs. 4B and 4D: By comparing the retention time, pregnenolone (1) seems to be eluted earlier than progesterone (2), which is supposed not to be such result using C18 column.
6) Lines 110, 112, 167 and 185: supporting information is not specified.
7) CYP204 catalyzes cholesterol and other phytosterols with lower efficiency compared to cholestenone. Please explain why. Can docking of cholesterol and other phytosterols into the active site give some evidence?
8) Cholesterol is much more wild-spread than cholestenone in organisms. Is it possible that cholesterol is not the direct substrate of CYP204 but catalyzed by another enzyme first to form cholestenone?
9) Please unify the format of cited references.

Reviewer #3

(Remarks to the Author)
This article by Tian et al. details the discovery and characterization of a series of bacterial P450scc enzymes that facilitate the side-chain cleavage of steroids in steroid catabolism. By conducting a comparative analysis of the phylogenetic origins of bacterial and mammalian P450scc enzymes, the authors highlight the convergent evolution of these isoforms. This study demonstrates the potential of bacterial P450scc enzymes as versatile and engineerable biocatalytic platforms for the synthesis of pregnenolone and progesterone, key precursors in steroid drug production. Through targeted mutagenesis, the authors show how specific amino acid substitutions influence substrate positioning and alter product profiles and catalytic activity, underscoring the complexity of enzyme dynamics that govern activity towards particular substrates. They further compare the mechanism and activity of the bacterial variants to their mammalian counterparts. Throughout this work, the authors utilize a comprehensive suite of techniques, including bioinformatics, biochemical assays, X-ray crystallography, and molecular simulations to support their claims. Overall, the manuscript is thorough and well-structured. The findings are well-justified, informative, and suitable for publication in Nature Communications after several minor revisions.

There are two minor points that need to be addressed prior to acceptance for publication of the manuscript.

1) Figure 4E: The arrow pushing needs to be fixed and it is suggested that the proposed mechanism is shown in more complete detail.

The arrow-pushing mechanism in Figure 4E is incorrect: The half arrows drawn from the porphyrin radical to the double bond of the Fe(IV)=O are in the wrong direction. Upon initiation of hydroxylation by Cpd-I, the porphyrin is reduced to neutral and Cpd-II, an Fe(IV)OH, is formed. In other words, one of the double bonds is homolytically cleaved with one electron going to the porphyrin and one electron going to form a bond with the substrate hydrogen that is abstracted during rebound. Additionally, the mechanistic scheme is somewhat confusing to look at since the outcome of the rebound mechanism is not drawn but instead cycle 3 with Cpd-I again is shown directly after the arrow. It would be helpful to see a complete version of what the authors are proposing.

2) The electron density for 9UIG should be reexamined.

While the structure of CYP203A3 (9UIG) is generally well-determined and the statistics are quite good, the electron density needs to be re-examined. There is electron density in several places in the map, but in particular in the F/G loop that is unattributed (i.e. there is significant residual positive Fo-Fc density) that has clear secondary structure and may represent the unmodeled components of the CYP204A3 structure. The authors made a reasonable choice to leave some parts of the protein unmodeled in the absence of certainty about how to connect the chains. However, examination of the electron density suggests that the model could be further improved especially near residues 181 – 192.

Minor suggestions:

1.Throughout the results section, the authors start many paragraphs with infinitives in a very repetitive fashion that could be changed to improve readability and flow of the manuscript:

Line 129: To evaluate
Line 176: To further determine
Line 188: In order to further investigate
Line 211: To elucidate the
Line 221: To investigate this
Line 245: To test this hypothesis,
Line 266: To further enhance
Line 293: For the purpose to elucidate the

2. In the sentence on Line 178: “The results showed that only 13 and 16 were further converted into 2 (Fig. 4B), whereas 14 and 15 were unreactive, confirming 13 and 16 as intermediates in the CYP204A5 catalytic process.”

The word “confirming” should be changed to “supporting” as this experiment reasonably supports the involvement of 13 and 16, but there are several alternative hypotheses that one could propose to achieve these results (e.g. changes in substrate binding/ reorientation). The findings from this experiment are well-justified, but “confirming” a mechanism or intermediates is too strong of language and would require much stronger (and challenging to obtain) evidence.

3. In Figure 2 Line 120: “Maxingmum Likelihood method.” Should be Maximum

4. In Figure 6: Only 1 of the porphyrin nitrogens is bonded to the Fe. All 4 nitrogens should be bonded (or none of them as in Figure 5).

5. In Figure 6: It would be more informative to plot B, D, and F as histograms to get a sense of the distributions and differences in distances between variants.

6. The molecular structures of 13, 16, 17 come after they are referenced in Figure 2. While they are introduced in the text, it would be helpful to introduce the structures prior to referencing them in a figure.

7. In Figure 5 Lines 258-259: “The distances between iron-oxo species of CYP204A3 and C20/22 of cholestenone are indicated with yellow dash lines”. There is no iron-oxo in this figure. The distances appear to be between the Fe and carbons directly.

Reviewer #4

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors have put in substantial effort into revising their manuscript, addressing most of our concerns, in part through conducting additional experiments and simulations.
However, several critical points remain to be revised before we can recommend this article for publication. Note that this review is mainly focused on the molecular modeling and simulation components of the manuscript.

The most important ones are:

-While the authors have now included some raw simulation data (input/output pdbs, some box sizes and some non-bonded force field parameters), the data currently provided would not be sufficient to reproduce their results. Full MD-trajectories (.xtc/.trr files with GROMACS) are likely too big to be uploaded but it would be necessary to include all other files (.mdp, .top, .gro, .log, .ndx,...) in order to understand and recreate their results.
-Likewise, all docking poses should be included.
-The Gaussian input & output files are also missing.

-The newly added comparison between the Alphafold model of CYP204A3 & their experimental structure (Supplementary Fig. S9), would profit from a close-up of the binding-site residues (side-chain representations). There is a notable deviation in the variable F-G loop region which could affect ligand binding and this should be stated in the main text where the alphafold model is mentioned.

-In addition to the ligand-RMSDs reported for the new MD-simulations of CYP204A3:substrate vs. CYP11A1-Cholesterol, the claim that CYP204A3 displays higher active site plasticity cannot be assessed merely by ligand RMSD. Additional RMSD measurements of the active-site residues would be necessary to substantiate this claim.

-In the response to our question about the docking scores, it is not clear if -8.976 kcal/mol is the score of the best or second best pose (used for modeling). Please also include all docking scores in a Supplementary table, along with other criteria that were used to select the relevant pose.
-In the same section, the authors reference the newly added Supplementary Fig. S11. From the caption of this figure, it is not stated which kind of data is presented. The caption should be revised to specify the kind of spectra.

-In their response to our questions concerning the SwissDock parameters, the authors refer to the newly added Supplementary Table S6. This table does not include any information on SwissDock. However, it reports the usage of Autodock Vina and another docking method named ‘AC’.
Neither Autodock Vina nor the other program is mentioned anywhere in the main text.
Which docking program did the authors use? If Vina was used, it should be described and referenced. Also, it is unclear to us what ‘AC’ refers to. We are not familiar with any program of this name. It should be discussed & referenced correctly. If ‘AC’ is an abbreviation, the full name should be given. Table S6 also lacks information on units.

-In their response to our concerns on spin-contamination, the authors correctly conclude that their reported <S^2> value is low enough to merit a very accurate DFT calculation. However, they only report this value in their response to our comments. Please include it in the Supplementary material so that future readers also know that these computations are reliable. Likewise, the imposed spin-multiplicity should be included. Both of these concerns can be addressed by making the Gaussian input and output files available.

-The authors should report the precise procedure utilized to derive RESP with Multiwfn charges in the Supplementary Material. The program offers many options during this procedure e.g. in the inclusion of constraints, the choice of fitting points and many more.

-Figure 4E is still very hard to read. It is not immediately clear that each reaction step involves an entire catalytic cycle of the P450 and the inclusion of the ferric, water-bound resting state at the bottom of each figure part makes it even more confusing. As the reactions are also depicted in Fig. 1(C), it might be easier to entirely remove this panel (4E) and explicitly mention that the enzymes catalyze two reactions which both involve Cpd1 as the active species.

-In Figure 6, histograms are given of distances analysed for the wild-type and mutant CYP 204A5. The authors need to provide information on what the curves are in this figure and the corresponding figures in the SI. If they are Gaussian fits, are these justified as they do not always fit the data well, especially for distance d2? A better treatment might omit the fit lines

and only consider the position and height of the highest histogram bars. This would still support the trends towards shorter distances in the mutants.

Reviewer #2

(Remarks to the Author)
All of my comments have been addressed comprehensively. I recommend publication in Nature Communications.

Reviewer #3

(Remarks to the Author)
The authors have made a significant effort to address my (and other reviewers') concerns and sufficiently addressed a majority of the comments raised.

Reviewer #4

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
While the authors have addressed all our comments on the previous versions in their rebuttal, the accompanying edits to the manuscript are unfortunately not complete and require revision as follows:

-GROMACS files: While the authors provide files for production runs, the input files for energy minimization and equilibration still appear to be missing.

-RESP: While the files used seem to be present in the supplementary dataset, their usage is not really clear from reading step.txt and, as the procedure used is not standard, the description given in the responses helps to clarify what was done and should be incorporated in the manuscript.

-Docking: The methods used have been clarified. However, Table S6 only includes selected docking scores rather than all docking scores. E.g. the score of -8.976 kcal/mol is for the second best score (according to the response) but the value for the best score is not given. In addition, its not clear from the table that this is the second best score and what the ranking of the other scores is. As docking scores always have to be considered in context, the complete output pose and scoring information should be given. The additional criteria in the footnote are qualitative: what are the quantitative criteria for accepting a pose?

Reviewer #5

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

The study by Tian et al. present an impressive effort in identifying biotechnologically bacterial cytochrome P450 enzymes that catalyse steroid side-chain cleavage by employing a plethora of biochemical, bioinformatic & biophysical methods.

However, the manuscript lacks some important information, and has presentational issues and draws conclusions from incomplete analyses.

Upon revision, the points given below should be addressed.

Summary:

The study aims at identifying bacterial C14-alpha hydroxylase Cytochrome P450 (CYP) enzymes for biotechnological applications but unexpectedly revealed a bacterial enzyme catalysing steroid side-chain cleavage. These bacterial CYPs bear the advantage of solubility when compared to their membrane-bound mammalian counterparts.

To achieve this, the authors applied an impressive, holistic approach encompassing identification of candidate enzymes by bioinformatic- & biochemical screening, investigation of substrate-specificity & enzymatic mechanism by structural biology, molecular dynamics simulations and biochemical assays & enhancement of enzymatic activity through mutagenesis experiments.

The remarkably wide scope of methods & characterization approaches presented in this manuscript necessitates a wide range of areas of expertise for an appropriate review. This review will, thus, mostly focus on the computational- (docking & MD simulation) and structural sections of the manuscript.

Initially, candidate enzymes were identified from bioinformatic screening and were subjected to experimental screening on different Steroid substrates. Only one catalytically active enzyme: P450-A8 was found to transform cholestenone to Progesterone while also working on a wider substrate range (different from eukaryotic equivalents). This sequence was used to identify 8 further homologous bacterial enzymes for further testing which revealed CYP204A5 to have the highest activity.

To solve the enzymatic mechanism, mono- & dihydroxylated intermediates were identified through comparative LC-MS & chemical synthesis indicating a 2-step hydroxylation mechanism.

This was further supported by structural studies on CYP204A3 (A5 was not crystallizable) & associated docking and molecular dynamics simulations. From the latter and site-saturation mutagenesis experiments, residues facilitating substrate positioning &

enzymatic activity were identified yielding novel mutants with increased activity. Further MD studies on 3 representative mutants demonstrated that these mutations place the C20 & C22 atoms closer to the catalytically active Compound 1 species.

Response:

We sincerely appreciate your insightful and constructive feedback on our work, particularly the valuable comments and suggestions regarding the MD simulations and molecular docking analyses. These comments have significantly strengthened the rigor and interpretation of our computational findings. In response, we have implemented the following key revisions:

(1) Revised all MD and docking protocols with enhanced parameters and validation steps.; (2) Conducted three independent MD replicates to ensure reproducibility; (3) Added quantitative comparisons (e.g., RMSD); (4) Performed complementary mutational biochemical assays to corroborate computational predictions.

On the other hand, despite extensive efforts, we were unable to obtain co-crystal structures of the protein with its steroid substrates. Given CYP204-mediated multi-step catalytic mechanism, molecular docking and MD simulations alone cannot fully elucidate the process. We therefore interpret the MD simulation results as a possible mechanism that may contribute to the observed activity enhancement. Our core conclusions regarding the novel activity and mechanism of CYP204A remain robust.

The revised manuscript now provides a more balanced interpretation of these computational results while maintaining the strength of our experimental findings.

The point-by-point response is below.

Upon revision, the points given below should be addressed.

Major concerns:

-It is unclear if CYP204A3, whose structure was solved crystallographically, differs from the most efficient candidate identified in this study, CYP204A5, in substrate specificity. The authors only remark on its lower activity, but a thorough comparison of enzyme specificity would be necessary to assess how good of a template this structure is for CYP204A5. Likewise, some measures like sequence identity and sequence alignment would should be provided to assist assessment of the authors′ choice to crystallize this enzyme.

Response:

We thank the reviewer for this important point. We supplemented these following data demonstrate that despite CYP204A3's lower catalytic activity it shares key functional

characteristics with CYP204A5:

(1) 70.3% sequence identity and conserved pocket residues (Supplementary Fig. S15).

(2) Identical substrate promiscuity profiles and intermediate production (Fig. 3A and Supplementary Fig. S8).

The reliability of our structural modeling approach is validated by the exceptional agreement between AlphaFold3's CYP204A3 prediction and our experimental structure (RMSD = 0.488 Å; Supplementary Fig. S9), demonstrating AlphaFold3's reliability for modeling the CYP204 family. While we extensively attempted CYP204A5 crystallization, its AlphaFold3 model proved reliable for our mechanistic studies. These data collectively support CYP204A3 as a valid structural proxy for CYP204A5."

We add these results now included in the revised manuscript Text-line237-246- clean version.

-From their comparison of the crystallographic structure of CYP204A3 to mammalian CYP11A1 the authors conclude that the former is less specific due to a looser fit of the docked substrate. This comparison seems very challenging as they compare between a crystallographic holo- and an apo-structure, thus, potentially missing out on induced-fit effects taking place upon substrate binding. It might be necessary to compare to the converged MD-structure of the docked protein-substrate complex, compare quantities like ligand RMSDs or even try to co-crystallize a steroid bound complex to substantiate this point.

Response: We thank the reviewer for raising this important point regarding the comparisons between CYP204A3 and CYP11A1.

We performed comparative MD simulations of substrate-bound CYP204A3 and the mammalian CYP11A1-cholesterol complex (PDB: 3N9Y). Quantitative analysis revealed higher ligand RMSD values in CYP204A3 (relative to protein backbone atoms; Supplementary Fig. S12), indicating greater conformational flexibility of the bound substrate. This enhanced active site plasticity may contribute to CYP204A3's broader substrate specificity compared to the more constrained CYP11A1.

In addition, we also systematically attempted to obtain co-crystal structures with multiple sterols, only the apo form of CYP204A3 was successfully determined. Given the complex nature of P450-substrate interactions, we interpret these findings cautiously and present them as one potential factor among several that may influence substrate range. We add these results and discussion included in the revised manuscript Text-line254-258-clean

version.

-The manuscript lacks any thorough validation of the docking poses of cholestenone to CYP204A3: what were the docking scores?

Response: We thank the reviewer for this valuable comment. The docking score for CYP204A3-cholestenone is predicted to be -8.976 kcal/mol. Notably, the highest-scoring conformation was not selected—instead the second. We prioritized: (1) Steric feasibility for C20/C22 of the side chain in close proximity to the heme center.; (2) refer to mammalian P450scc structural and mechanism analysis data.

To further validate the docking poses, we supplemented experimental evidence from alanine-scanning mutagenesis in CYP204A3 (Supplementary Fig. S11), which confirmed key active-site residues; Furthermore, motif-swapping experiments involving CYP204A3 and CYP204A5 provide strong support for the reliability of our chosen conformations (Fig. 5E).

We added related results and discussion included in the revised manuscript, Text-line 250-253- clean version.

How do the poses e.g. compare to alignments with CYP11A1:Cholesterole? Likewise, it would be good to give a proper interaction plot, not only the Surface-fit presented in Fig S8 and to provide the docked structure as supplementary material.

Response: We thank the reviewer for these constructive suggestions. We have now included a detailed structural alignment comparison between our CYP204A3-cholesterol docked complex and the CYP11A1-cholesterol crystal structure in Supplementary Fig. S10; The docked structure files have been provided as supplementary Data: MD-simulation-Data Dataset 3.

-Likewise, any critical validation of the performed MD simulation is missing. Without any measure of convergence besides the C20, C22-Fe distances, it is impossible to assess if the simulation-lengths & choice to only simulate one replica per system are sufficient. This is very important as some of these distance-plots e.g. in Fig. 6 B & D as well as S11 B & S12 B display increases/instabilities just before the simulations were terminated indicating some structural rearrangements.

Such measures could include e.g. protein & ligand RMSDs over all trajectories.

Response: We appreciate the reviewer's insightful comment regarding the validation of

simulation convergence. To rigorously address this concern, we have now performed three independent replicas for each MD simulation and analyzed the backbone RMSD of all trajectories to ensure system stability (Supplementary Fig. S12, S13, S17, S19 and S20).

-In the sentence ‘Closer examination of the substrate-binding pocket further identified several key amino acid residues […]’ (p. 9) it is unclear how this examination was performed. Was this based on quantitative measurements such as contact analyses or on visual inspection? This information is necessary to assess the validity of all subsequent experiments.

Response: We thank the reviewer for this important clarification. The key residues were identified through analysis of docking models using PyMOL (interaction distance < 5 Å). We have revised the text to state: " Analysis of the docked model reveals potential ligand–receptor interactions mediated by several key amino acid residues within 5 Å—specifically S80, M81, F84, Y93, S94, L253, W256, A257, and T261—which may play a pivotal role in positioning the side chain of **CHO** above the heme group" in Text-line 260-263-clean version to clearly reflect the analytical approach.

-The ‘Methods’ & ‘Supplementary methods’ sections lack essential information on the simulation-based assays:

-Which parameters were used in SwissDock ?

Response: We adjusted the position and size of the box to align it above the heme of P450 and meet the online computational limitation. Other system's default parameters for the rest. The relevant parameters are added in Supplementary Table S6.

-Which implicit solvent model was used in DFT?

Response: The solvent effects were modeled using the integral equation formalism variant of the polarizable continuum model (IEFPCM). We have now explicitly stated this in the revised Methods section: ″To account for solvent effects, the IEFPCM implicit solvation model was applied throughout the optimization [Chemical Physics 55.1 (1981): 117-129. & Chemical reviews 105.8 (2005): 2999-3094.].″.

-Which spin-state was chosen to represent CPD1? How high was the associated spin-contamination? High spin-contamination can over-delocalize charges over the molecule,

thus, affecting RESP calculations.

Response: A quartet spin-state (spin multiplicity = 4) was used for our model of Compound I (CPD1). This choice is well-precedented and consistent with previous DFT studies on the heme Compound I intermediate (e.g., J. Comput. Chem. 2012, 33 (2), 119). To assess the degree of spin contamination, we analyzed the expectation value of the spin-squared operator, $\langle S^2 \rangle$. The calculated $\langle S^2 \rangle$ value before annihilation was 3.7880. This value is very close to the ideal theoretical value of 3.75 for a pure quartet state (S=1.5). The deviation is only ~1.01% from the ideal value, indicating a very low and acceptable level of spin contamination. Furthermore, the $\langle S^2 \rangle$ value after the annihilation of the first spin contaminant is 3.7507, which is nearly identical to the ideal value, further confirming the appropriateness of the chosen wavefunction. We agree with the reviewer that high spin-contamination can be a concern for charge distribution. However, given the minimal spin contamination observed in our system, we are confident that its effect on charge delocalization and the reliability of the subsequent RESP calculations is negligible. We have added the following clarification to the Methods section: ″The spin multiplicity was set to 4, in agreement with previously published studies [J. Comput. Chem. 2012, 33 (2), 119–133].″.

-How were the RESP charges computed in Multiwfn? Which fitting algorithm etc. was utilized? How were the charges made compatible with AMBER?

Response: Similar to Antechamber, Multiwfn also applies the standard RESP method for deriving atomic charges, as originally described by Kollman, P. A. et al. (J. Phys. Chem. 1993, 97(40), 10269). That is, a two-step charge fitting is performed. In the first stage, a weak hyperbolic restraint is imposed on the atomic charges to allow the greatest degree of freedom for the charges on polar atoms to fit the ESP. In the second stage, the fitting is restricted to the charges of sp3 hybridized carbons, methylene carbons, and their attached hydrogen atoms, while enforcing charge equivalency for chemically equivalent hydrogens. Further details of this procedure are provided in section 3.9.16 of the Multiwfn software manual. The application of Multiwfn for generating RESP charges is well-established and has been reported in numerous publications (e.g., Nat. Commun., 2024, 15 (1), 6249).

why did the authors choose to use Multiwfn for RESP charge computations instead of the native AMBER program Antechamber which can also compute RESP charges from ESP

calculations with Gaussian. Likewise, why was the ′sobtob′ tool utilized instead of Antechamber

Response: The parameterization of the iron-containing complex presented a challenge, as the standard Antechamber protocol lacks support for RESP charge derivation involving iron centers. To circumvent this limitation, we employed Multiwfn for the charge calculation. Furthermore, while AmberTools is integral to the Amber simulation package, its direct application here was impractical due to inherent format incompatibilities with GROMACS, our chosen MD engine. We therefore leveraged the Sobtop program to construct the molecular topology. Sobtop was uniquely suited for this task as it systematically assigns GAFF atom types where applicable, while seamlessly integrating parameters from the Universal Force Field (UFF) for atoms not covered by GAFF, such as the central iron ion. Sobtop generates GROMACS-compatible topology files that maintain consistency with the overarching Amber force field, providing a robust and streamlined workflow for our simulations.

How was the system protonated and at what pH?

Response: We performed protonation at pH 7.4 using gmx pdb2gmx (GROMACS) with the force-field default protonation rules. Asp/Glu were set deprotonated, Lys/Arg protonated, and termini in their zwitterionic forms. Histidines were assigned as neutral HID/HIE tautomers according to local hydrogen-bonding suggested by pdb2gmx. We have added these details to the Methods section.

References for the AMBER14SB-PARMBSC1 and GAFF force fields and the TIP3P water model should be given.

Response: Related references have been added to the supplementary method.

AMBERFF14SB：Maier, J.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB J. Chem. Theory Comput. 2015, 11, 3696-3713.

GAFF force fields：Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A. (2004), Development and testing of a general amber force field. J. Comput. Chem., 25: 1157-1174. https://doi.org/10.1002/jcc.20035

TIP3 water model：William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, Michael L. Klein; Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 15 July 1983; 79 (2): 926 – 935.

https://doi.org/10.1063/1.445869

-How many counter-ions were added to the simulation systems, i.e. at which concentration?

Response: We have clarified this in the Methods section with the following addition: " Four sodium ions were added to neutralize the total charge of the system"; no additional ions were included."

-No information on electrostatics, treatment of periodic boundary condition etc. are given for the energy minimizations

Response: - We are grateful to the reviewer for their diligent feedback, which has been instrumental in helping us identify and address the descriptive shortcomings within the Methods section. We have now specified the initial periodic boundary conditions for the system in the Methods section: ″The protein complex was solvated in a rectangular box, with periodic boundary conditions applied in all directions. The exact dimensions for each simulation system are provided in checklist-4a.xlsx.″

Furthermore, the description of the electrostatic interactions has been revised to clarify its application throughout all simulation stages, including energy minimization: ″Throughout the energy minimization and subsequent simulation steps, long-range electrostatic interactions were treated with the particle mesh Ewald (PME) method using a 1.2 nm cutoff for the Coulomb interactions. Van der Waals (vdW) interactions cutoff was also set to 1.2 nm″

-The ′Methods′ section of the manuscript essentially only describes experimental methods. Despite majorly contributing to the manuscript, almost all bioinformatics- and simulation-related methods have to be searched for in the Supplementary materials. Especially in the light of our concerns about critical information lacking there, it is unclear to us why the methodology was fragmented like this.

Response: We appreciate the reviewer's valid concern regarding the organization of methodological details. To improve clarity and accessibility, we have restructured the Methods section to integrate all computational and bioinformatics approaches alongside experimental procedures in the main manuscript.

Minor concerns:

-Language: some incomprehensible sentences, some repetitions/tautologies (e.g. ′[...] plant CYP87A family P450scc enzymes [...] have recently been identified in several plants′); many titles of Supplementary materials erroneous

Response: We have revised ambiguous sentences, removed tautologies (e.g., simplified CYP87A description), and corrected all supplementary material titles.

-The authors′ initial screening approach seems to undervalue the importance of local vs. global similarity. The way we understand their initial bioinformatic approach, it was based essentially on just selecting one template enzyme and a global alignment to identify initial hits. Since CYPs are very well known & relatively conserved, it might have been more successful if the authors would have instead focused on local alignments in the known substrate-coordinating regions of steroid-metabolizing CYPs. This point should be elaborated on in the introduction and/or discussion.

Response: We sincerely appreciate the reviewer's insightful suggestion regarding our bioinformatic approach. Our primary objective was to identify potential bacterial steroid P450 hydroxylases based on phylogenetic analysis with known steroid-metabolizing CYPs, as described in the initial mining section (Text-line 107-109).

We recognize that detailed analysis of substrate-coordinating regions would provide a more precise method for identifying steroid-specific P450s. We expanded the discussion section in Text-line 369-375-clean version to highlight the potential of future large-scale analyses examining sequence-structure-substrate relationships of known steroid-metabolizing CYPs. This structural bioinformatics approach could enable more targeted mining of steroid-metabolizing enzymes in future studies.

-The manuscript's clarity would greatly profit from using proper abbreviations instead of numbers to refer to different compounds e.g. PROG/PREG/..., particularly as this gets very confusing when reaction steps are discussed.

Response: We appreciate this excellent suggestion. In the revised manuscript, we have implemented the following changes: Systematic abbreviation use: Introduced standard steroid and intermediates abbreviations (e.g., PROG for progesterone, PREG for pregnenolone) at first mention.

-The text in Fig. 2 is too small and not readable

Response: The text in Figure 2 has been enlarged and reformatted for clarity.

-Fig. 4 E: the presented mechanism does not seem to make sense: where does the third oxygen in ′Cycle 3′ come from? Is it derived from a water molecule as in Cycle 3, CPD1 still carries an oxygen?

Response: The third oxygen in Cycle 3 originates from molecular oxygen (O_2). As depicted in the figure, the catalytic side-chain cleavage of P450scc is a stepwise oxidation reaction. Each oxidation step represents a new cycle of P450 activity. This process uses the classical pathway of P450 to form CpdI.

In accordance with Reviewer 3's suggestion and to enhance clarity, we have revised Figure 4E.

-Actual means & standard deviations for the distances measured over the MD simulation should be reported instead of discussing ′approximate values′ in paragraph 2 on p. 9

Response: We have revised in text as exact mean ± SD distances (4.04±0.41 Å and 4.10±0.45 Å, respectively) are now reported.

-The title of table S2 is entirely unclear to us (′Information of redox used in this study′)

Response: The title has been revised to: Redox protein systems utilized for P450 activity in this study.

-In the Supplementary methods, the authors refer to AF3 models as homology models which is wrong.

Response: We corrected "homology models" to " The predicted structures" in Supplementary Methods.

-Why did the authors choose the AMBER14SB-PARMBSC1 force field?

Response: We acknowledge the potential for confusion and have removed the reference to PARMBSC1, which is a force field specifically parameterized for DNA and is not relevant to our work. Our study exclusively employed the AMBER14SB force field, a choice motivated by its demonstrated robustness and accuracy in protein simulations. The suitability of this force field is further substantiated by its successful application in recent, comparable studies (e.g., JACS Au 2024, 4 (4), 1591 – 1604).

Reviewer #2 (Remarks to the Author):

The manuscript by Tian et al. describes the discovery of bacterial P450scc enzymes in steroid-degrading microorganisms. The authors reveal that these bacterial enzymes proceed via a different mechanism compared to the mechanism of the comparable enzymes found in mammals and plants. The authors also perform engineering work based on one obtained bacterial P450scc enzyme structure. Overall this is an excellent paper, providing new insights into bacterial metabolism and also highighting a new biocatalytic tool for production of progesterone-derived compounds.

Response： We thank the reviewer for high appreciation of this work.

Below are minor commenst:

1) The stereochemistry of the compounds in Fig. 1 is neither complete (1, 2 ,3, and 6), nor correct (Fig. 1C, intermediate 6).

Response： The stereochemistry of compounds has been corrected and completed in the revised Fig.1.

2) The intermediates in Fig. 1. are numbered wrongly.

Response：The numbering in Fig. 1 has been corrected

3) Please make Fig. 3B more reader friendly since some annotations are upside-down.

Response：Annotations have been reoriented for clarity in the revised figure

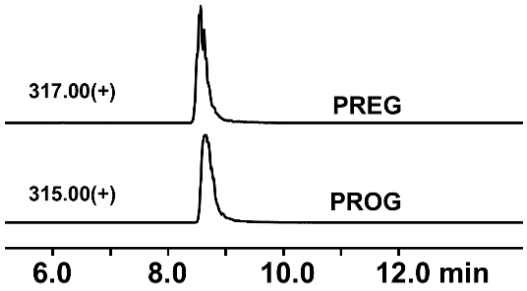
4) Please explain why the peaks for **5** in Fig. 4D are not consistent by retention time.

Response： As described in the main text (in Text-line 212-213 clean version), these peaks with MW 419 in Fig 4D CYP204A5+ 20S-HC represent unidentified dihydroxylation CHL products, which are not **5** (20R,22R-DHC). To avoid confusion, they have been annotated in the newly revised legend and figure.

5) Figs. 4B and 4D: By comparing the retention time, pregnenolone (1) seems to be

eluted earlier than progesterone (2), which is supposed not to be such result using C18 column.

Response : Thanks for your careful examination. This deviation is primarily due to the fact that Figure 4B presents an EIC from UPLC-MS analysis, in which the retention times exhibit a slight delay compared to the corresponding UPLC chromatogram. The ion chromatograms of PREG and PROG are provided in the figure below to more clearly illustrate the difference in retention times between the two compounds. In fact, these two compounds have very similar retention times in our analysis.



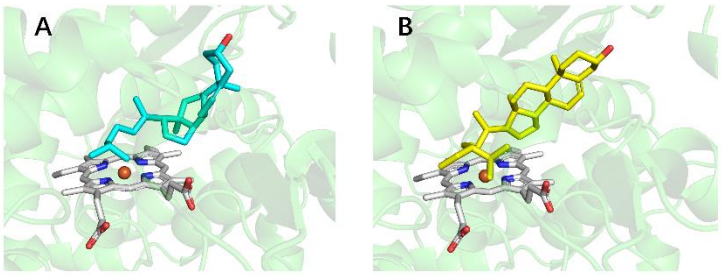
6) Lines 110, 112, 167 and 185: supporting information is not specified.

Response : Lines 110, 112, 167, and 185 refer to NMR structural data. To improve clarity and maintain consistent numbering of figures between the main text and the supplementary material, we have retitled this section in main text as ″Supplementary spectral Data.″ The supporting information file includes complete spectral details for each relevant compound.

7) CYP204 catalyzes cholesterol and other phytosterols with lower efficiency compared to cholestenone. Please explain why. Can docking of cholesterol and other phytosterols into the active site give some evidence?

Response : Docking results of CYP204A3 with cholesterol (A) and other phytosterols (B) are provided below. Combined with MD simulation analysis for cholestenone (CHO), we propose that the lower catalytic efficiency of CYP204 toward cholesterol and phytosterols stems from differences in substrate-enzyme interactions: (1) The 3-keto group in cholestenone forms distinct hydrogen bonds with active site residues, unlike the 3-hydroxyl group in cholesterol; (2) Phytosterols experience steric hindrance due to their varied alkyl side chains, in addition to the 3-hydroxyl group, which disrupts optimal

positioning of C20/C22 near CpdI. We add related discussion in Text-line 249-254 clean version.



8) Cholesterol is much more wild-spread than cholestenone in organisms. Is it possible that cholesterol is not the direct substrate of CYP204 but catalyzed by another enzyme first to form cholestenone?

Response：This is indeed plausible, as 3β-hydroxysteroid dehydrogenases (3β-HSDs) are widely present in steroid-metabolizing bacteria （Microorganisms 2021, 9, 469. https://doi.org/10.3390/microorganisms9030469）and could catalyze the oxidation of cholesterol's 3β-hydroxyl group to form cholestenone prior to CYP204-mediated modification.

9) Please unify the format of cited references.

Response：we have unified the format of cited references.

Reviewer #3 (Remarks to the Author):

This article by Tian et al. details the discovery and characterization of a series of bacterial P450scc enzymes that facilitate the side-chain cleavage of steroids in steroid catabolism. By conducting a comparative analysis of the phylogenetic origins of bacterial and mammalian P450scc enzymes, the authors highlight the convergent evolution of these isoforms. This study demonstrates the potential of bacterial P450scc enzymes as versatile and engineerable biocatalytic platforms for the synthesis of pregnenolone and progesterone, key precursors in steroid drug production. Through targeted mutagenesis, the authors show how specific amino acid substitutions influence substrate positioning and alter product profiles and catalytic activity, underscoring the complexity of enzyme dynamics that govern activity towards particular substrates. They further compare the mechanism and activity of the bacterial variants to their mammalian counterparts. Throughout this work, the authors utilize a comprehensive suite of techniques, including bioinformatics, biochemical assays, X-ray crystallography, and molecular simulations to support their claims. Overall, the manuscript is thorough and well-structured. The findings are well-justified, informative, and suitable for publication in Nature Communications after several minor revisions.

Response: We are sincerely grateful to the Reviewer for their generous appreciation and positive assessment of our study.

There are two minor points that need to be addressed prior to acceptance for publication of the manuscript.

1) Figure 4E: The arrow pushing needs to be fixed and it is suggested that the proposed mechanism is shown in more complete detail.

The arrow-pushing mechanism in Figure 4E is incorrect: The half arrows drawn from the porphyrin radical to the double bond of the Fe(IV)=O are in the wrong direction. Upon initiation of hydroxylation by Cpd-I, the porphyrin is reduced to neutral and Cpd-II, an Fe(IV)OH, is formed. In other words, one of the double bonds is homolytically cleaved with one electron going to the porphyrin and one electron going to form a bond with the substrate hydrogen that is abstracted during rebound. Additionally, the mechanistic scheme is somewhat confusing to look at since the outcome of the rebound mechanism

is not drawn but instead cycle 3 with Cpd-I again is shown directly after the arrow. It would be helpful to see a complete version of what the authors are proposing.

Response：Thanks for your suggestions, the Figure 4E have been revised for clarify its mechanism.

2) The electron density for 9UIG should be reexamined.

While the structure of CYP203A3 (9UIG) is generally well-determined and the statistics are quite good, the electron density needs to be re-examined. There is electron density in several places in the map, but in particular in the F/G loop that is unattributed (i.e. there is significant residual positive Fo-Fc density) that has clear secondary structure and may represent the unmodeled components of the CYP204A3 structure. The authors made a reasonable choice to leave some parts of the protein unmodeled in the absence of certainty about how to connect the chains. However, examination of the electron density suggests that the model could be further improved especially near residues 181 – 192.

Response： We thank the reviewer for highlighting the need to re-examine the electron density in the CYP203A3 (9UIG) structure, particularly in the F/G loop region (residues 181-192). In response, we have carefully reanalyzed the electron density maps and successfully modeled all previously unassigned residues in this region. The updated model shows good agreement with the electron density, though some flexibility remains evident in parts of the loop. We have addressed the residual Fo-Fc density by properly incorporating these residues into the refined structure. To support our revisions, we provide the updated PDB (9WAT) and MTZ files along with a figure showing the 2Fo-Fc electron density (contoured at 1.0 σ) for residues 182-192, demonstrating the improved model quality. These modifications have enhanced both the completeness and accuracy of the structural model. We add revised file in Supplementary attachment: CYP204A3 structure data for reviewers Dataset 4.

Minor suggestions:

1.Throughout the results section, the authors start many paragraphs with infinitives in a very repetitive fashion that could be changed to improve readability and flow of the manuscript:

Line 129: To evaluate

Line 176: To further determine

Line 188: In order to further investigate

Line 211: To elucidate the

Line 221: To investigate this

Line 245: To test this hypothesis,

Line 266: To further enhance

Line 293: For the purpose to elucidate the

Response : We revised the manuscript to improve readability by varying sentence structure throughout the Results section, eliminating the repetitive use of infinitives to begin paragraphs.

2. In the sentence on Line 178: ″The results showed that only 13 and 16 were further converted into 2 (Fig. 4B), whereas 14 and 15 were unreactive, confirming 13 and 16 as intermediates in the CYP204A5 catalytic process.″The word ″confirming″ should be changed to ″supporting″ as this experiment reasonably supports the involvement of 13 and 16, but there are several alternative hypotheses that one could propose to achieve these results (e.g. changes in substrate binding/ reorientation). The findings from this experiment are well-justified, but ″confirming″ a mechanism or intermediates is too strong of language and would require much stronger (and challenging to obtain) evidence.

Response : We changed "confirming" to "supporting" to more accurately reflect the level of evidence provided by these experimental results regarding intermediates.

3. In Figure 2 Line 120: ″Maxingmum Likelihood method.″ Should be Maximum

Response : The misspelling "Maxingmum" has been corrected to "Maximum" in Figure 2.

4. In Figure 6: Only 1 of the porphyrin nitrogens is bonded to the Fe. All 4 nitrogens should be bonded (or none of them as in Figure 5).

Response : We have corrected the iron-porphyrin coordination in Figure 6 to properly show all four nitrogen atoms bonded to the central iron.

5. In Figure 6: It would be more informative to plot B, D, and F as histograms to get a sense of the distributions and differences in distances between variants.

Response : We appreciate this excellent suggestion. The distance distribution data in Figure 6 have been revised to better visualize the statistical distributions and differences

between variants.

6. The molecular structures of 13, 16, 17 come after they are referenced in Figure 2. While they are introduced in the text, it would be helpful to introduce the structures prior to referencing them in a figure.

Response : We have replaced numerical designations with standardized abbreviations for compounds and intermediates, as suggested by the reviewers 1, to enhance clarity and differentiation throughout the text.

7. In Figure 5 Lines 258-259: ″The distances between iron-oxo species of CYP204A3 and C20/22 of cholestenone are indicated with yellow dash lines″. There is no iron-oxo in this figure. The distances appear to be between the Fe and carbons directly.

Response : We have revised the description of Figure 5 to accurately reflect that the measured distances are between the iron center and substrate carbons (C20/22), not an iron-oxo species.

Reviewer #4 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Response: We sincerely appreciate the constructive comments and the time dedicated to evaluating our work.

Reviewer #5 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Response: We sincerely appreciate the constructive comments and the time dedicated to evaluating our work.

Reviewer #1 (Remarks to the Author):

The authors have put in substantial effort into revising their manuscript, addressing most of our concerns, in part through conducting additional experiments and simulations.

However, several critical points remain to be revised before we can recommend this article for publication. Note that this review is mainly focused on the molecular modeling and simulation components of the manuscript.

The most important ones are:

-While the authors have now included some raw simulation data (input/output pdbs, some box sizes and some non-bonded force field parameters), the data currently provided would not be sufficient to reproduce their results. Full MD-trajectories (.xtc/.trr files with GROMACS) are likely too big to be uploaded but it would be necessary to include all other files (.mdp, .top, .gro, .log, .ndx,...) in order to understand and recreate their results.

-Likewise, all docking poses should be included.

-The Gaussian input & output files are also missing.

Response: We sincerely thank the reviewer for this important comment concerning reproducibility. In response, we have now provided all essential GROMACS input and output files, docking poses and Gaussian input & output files in supplementary database.

For GROMACS input and output files, the .tpr files were generated using GROMACS version 2022.4, and they contain complete information on the system topology, force field parameters, and simulation settings. These files can be directly executed with the command gmx mdrun -s filename.tpr, allowing other researchers to fully reproduce the reported simulations. Due to the large size of the .xtc (trajectory) and .log (run log) files —over several gigabytes in total across all replicates—these files were not uploaded. However, the provided files include all necessary information for verification and reproduction of the simulations, ensuring transparency and reproducibility of our computational results.

These data can be found in MD-simulation-Data Dataset 3.

-The newly added comparison between the Alphafold model of CYP204A3 & their

experimental structure (Supplementary Fig. S9), would profit from a close-up of the binding-site residues (side-chain representations). There is a notable deviation in the variable F-G loop region which could affect ligand binding and this should be stated in the main text where the alphafold model is mentioned.

Response: We thank the reviewer for this suggestion. As requested, we have added the binding-site residues to Supplementary Fig. S9B. We highlighted the deviation in the F-G loop region in the main text where the AlphaFold model is mentioned (line 233-234 in main text).

-In addition to the ligand-RMSDs reported for the new MD-simulations of CYP204A3: substrate vs. CYP11A1-Cholesterol, the claim that CYP204A3 displays higher active site plasticity cannot be assessed merely by ligand RMSD. Additional RMSD measurements of the active-site residues would be necessary to substantiate this claim.

Response: We appreciate your valuable suggestion. Following this advice, we calculated the RMSD values of the active-site residues for both CYP204A3 and CYP11A1 during the MD simulations. As shown in the newly added Supplementary Fig. S12C and D, the active-site residues of CYP204A3 exhibit higher RMSD values and broader fluctuations than those of CYP11A1, indicating greater conformational flexibility. Revisions to the main text are also included in line 243-244 in main text.

-In the response to our question about the docking scores, it is not clear if -8.976 kcal/mol is the score of the best or second best pose (used for modeling). Please also include all docking scores in a Supplementary table, along with other criteria that were used to select the relevant pose.

Response: Thank you for this important clarification. The docking score of −8.976 kcal/mol corresponds to the second-best pose, which was selected for further modeling based on its favorable binding conformation and interaction profile. For transparency, we have included all docking scores in Supplementary Table S6, along with detailed criteria used for pose selection.

-In the same section, the authors reference the newly added Supplementary Fig. S11. From the caption of this figure, it is not stated which kind of data is presented. The

caption should be revised to specify the kind of spectra.

Response: We appreciate the reviewer's careful attention to detail. We have revised the caption of Supplementary Fig. S11 to explicitly state that the figure presents. “UPLC profiles of CHO conversion by wild-type CYP204A13 and alanine mutants monitored at 254 nm.”

-In their response to our questions concerning the SwissDock parameters, the authors refer to the newly added Supplementary Table S6. This table does not include any information on SwissDock. However, it reports the usage of Autodock Vina and another docking method named ‘AC’.

Neither Autodock Vina nor the other program is mentioned anywhere in the main text. Which docking program did the authors use? If Vina was used, it should be described and referenced. Also, it is unclear to us what ‘AC’ refers to. We are not familiar with any program of this name. It should be discussed & referenced correctly. If ‘AC’ is an abbreviation, the full name should be given. Table S6 also lacks information on units.

Response: We appreciate the reviewer's insightful comments. Molecular docking was performed using the SwissDock web service (https://www.swissdock.ch/), which integrates two docking engines: Attracting Cavities (**AC**) and AutoDock Vina. Due to compatibility issues with the AlphaFold3-generated CYP204A5 structure in AutoDock Vina, the Attracting Cavities (**AC**) engine was employed. The selected docking pose from AC met all predefined evaluation criteria.

The revised manuscript now provides comprehensive methodological details in Supplementary Table S6, including parameter specifications and literature citations for both docking engines.

-In their response to our concerns on spin-contamination, the authors correctly conclude that their reported <S^2> value is low enough to merit a very accurate DFT calculation. However, they only report this value in their response to our comments. Please include it in the Supplementary material so that future readers also know that these computations are reliable. Likewise, the imposed spin-multiplicity should be included. Both of these concerns can be addressed by making the Gaussian input

and output files available.

Response: We thank the reviewer for this important observation. The Gaussian input and output files including S^2 value are now presented in MD-simulation-Data Dataset 3. We add this information in revised MD simulation method section.

Calculated < S^2> values for CpdI at B3LYP levels of theory

Method	Basis Set	Spin Multiplicity	<S^2> Calculated	<S^2> Ideal
B3LYP	6-31G(d)/Lanl2DZ	4	3.7880	3.75

-The authors should report the precise procedure utilized to derive RESP with Multiwfn charges in the Supplementary Material. The program offers many options during this procedure e.g. in the inclusion of constraints, the choice of fitting points and many more.

Response: Since the system is incompatible with Multiwfn's automated RESP fitting procedure (due to unsupported capping groups in the second step), we performed manual two-step ESP fitting following RESP principles. In the first step, we conducted initial ESP fitting with a weighting factor (a) of 0.0005 while constraining the total charges of formyl and amide groups to zero. The second step (a=0.0010) refined the charges by only permitting adjustments for sp^3 hybridized carbons, hydrogens (excluding capping groups), and two $=CH_2$ groups, with additional equivalence constraints applied to hydrogens in non-capping $-CH_3$, $-CH_2-$, and $=CH_2$ moieties.

The procedure is relatively complex; therefore, all files used for the RESP fitting are provided in the "HEME_4_4_RESP" directory. Detailed steps are described in step.txt, and the final charge results are given in HEME_4_4.chg. The files are added in MD-simulation-Data Dataset 3

-Figure 4E is still very hard to read. It is not immediately clear that each reaction step involves an entire catalytic cycle of the P450 and the inclusion of the ferric, water-bound resting state at the bottom of each figure part makes it even more

confusing. As the reactions are also depicted in Fig. 1(C), it might be easier to entirely remove this panel (4E) and explicitly mention that the enzymes catalyze two reactions which both involve Cpd1 as the active species.

Response: We thank the reviewer for this valuable suggestion. Upon careful consideration, we have removed the Figure 4E to improve clarity. Instead, we have clarified in the text that both enzymatic reactions utilize Cpd1 as the active species in line 205-210 in main text.

-In Figure 6, histograms are given of distances analysed for the wild-type and mutant CYP 204A5. The authors need to provide information on what the curves are in this figure and the corresponding figures in the SI. If they are Gaussian fits, are these justified as they do not always fit the data well, especially for distance d2? A better treatment might omit the fit lines and only consider the position and height of the highest histogram bars. This would still support the trends towards shorter distances in the mutants.

Response: We thank the reviewer for this suggestion. Following the recommendation, we have removed the Gaussian fit lines from Figure 6 and corresponding SI figures. The trends toward shorter distances in the mutants remain clearly evident from the histogram peak positions and heights.

Reviewer #1 (Remarks to the Author):

While the authors have addressed all our comments on the previous versions in their rebuttal, the accompanying edits to the manuscript are unfortunately not complete and require revision as follows:

-GROMACS files: While the authors provide files for production runs, the input files for energy minimization and equilibration still appear to be missing.

Response: We have now included the input files for energy minimization and equilibration in the updated Supplementary Data 10. (Figshare DOI: 10.6084/m9.figshare.30657758)

-RESP: While the files used seem to be present in the supplementary dataset, their usage is not really clear from reading step.txt and, as the procedure used is not standard, the description given in the responses helps to clarify what was done and should be incorporated in the manuscript.

Response: We have expanded the methodological description in the manuscript, incorporating the details provided in updated Supplementary Data 9. (Figshare DOI: 10.6084/m9.figshare.30657758)

-Docking: The methods used have been clarified. However, Table S6 only includes selected docking scores rather than all docking scores. E.g. the score of -8.976 kcal/mol is for the second best score (according to the response) but the value for the best score is not given. In addition, its not clear from the table that this is the second best score and what the ranking of the other scores is. As docking scores always have to be considered in context, the complete output pose and scoring information should be given. The additional criteria in the footnote are qualitative: what are the quantitative criteria for accepting a pose?

Response: We have included the output information for all docking pose results in the updated Supplementary Data 6. The selection criteria prioritize the option with the highest score that satisfies the first two criteria. This information has been added to the revised Supplementary Table 7.