JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

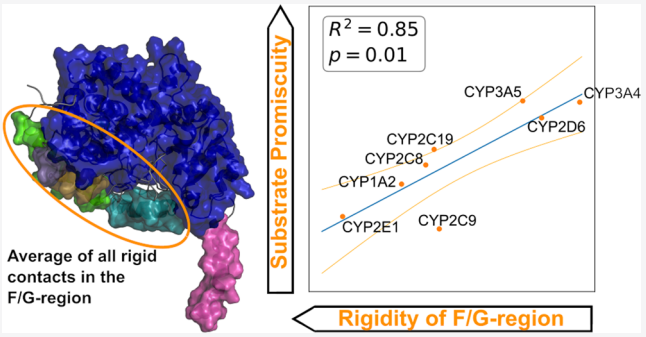
F/G Region Rigidity is Inversely Correlated to Substrate Promiscuity of Human CYP Isoforms Involved in Metabolism

Daniel Becker, Prasad V. Bharatam, and Holger Gohlke*

Cite This: J. Chem. Inf. Model. 2021, 61, 4023−4030 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Of 57 human cytochrome P450 (CYP) enzymes, 12 metabolize 90% of xenobiotics. To our knowledge, no study has addressed the relation between enzyme dynamics and substrate promiscuity for more than three CYPs. Here, we show by constraint dilution simulations with the Constraint Network Analysis for the 12 isoforms that structural rigidity of the F/G region is significantly inversely correlated to the enzymes' substrate promiscuity. This highlights the functional importance of structural dynamics of the substrate tunnel.



INTRODUCTION

Cytochrome P450 (CYP) enzymes play a central role in human drug metabolism.[1,2] They primarily act as monooxygenases, and the most essential catalyzed reaction is the hydroxylation of nonactivated C−H bonds.[3,4] This makes drugs and other xenobiotics more hydrophilic, which facilitates excretion, but also can inactivate drugs or, vice versa, lead to biologically active metabolites.[5] All human CYP enzymes share three common properties. First, CYPs are heme proteins with a similar fold. Second, the active site containing the heme is buried and only accessible via channels. Finally, human CYPs are generally associated with the membrane, usually the membrane of the endoplasmic reticulum, by an N-terminal anchor helix.[6,7] In total, 57 human CYP enzymes are known, but not all are involved in metabolism.[8,9] For some other CYP enzymes, the function is yet unknown.[10,11]

Interestingly, only 12 CYP isoforms metabolize about 90% of all xenobiotics.[1] These are CYP1A1, CYP1A2, CYP1B1, CYP2A6, CYP2B6, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP2E1, CYP3A4, and CYP3A5. Although all are substrate-promiscuous enzymes, they show recognizable differences, e.g., CYP3A4 contributes the most to xenobiotics metabolism by being involved in 20% of all known CYP transformations.[1] On the other hand, CYP1B1 contributes to only 3% of all known CYP transformations.[1] However, the sheer number of xenobiotics metabolism pathways a CYP contributes to is not the best indicator for its substrate scope. Substrate scope as a functional enzyme parameter should also take the enzyme's catalytic efficiency into account, e.g., although CYP3A5 generally catalyzes the same reactions as CYP3A4, it does so often at lower rates, such that CYP3A5 only contributes to 4% of all known CYP transformations.[1] Earlier, a molecular docking-based analysis pointed out that the degree of promiscuity may be partly determined by the amino acid residues in the dome region of the CYPs (CYP3A4—hydrophobicity-dominated, CYP2D6—acidic-dominated, CYP2C9—basicity-dominated residues).[12,13]

Previous work based on crystal structures and computational studies stressed the importance of the F/G region, consisting of the F, F′, G′, and G helices, for substrate binding[14−16] because the region is at the beginning of a tunnel that is considered the substrate entrance channel.[7] Such tunnels can act as filters and have been found to influence both substrate specificity and catalytic mechanism.[17−19] Accordingly, one study hypothesized for the three isoforms CYP3A4, CYP2C9, and CYP2A6 that higher mobility, deduced from movements of the C_α atoms, in the F/G region correlates with higher substrate promiscuity.[20] Enzyme structural dynamics, in addition to its role in catalysis[21,22] and allosteric regulation,[23−26] has also been recognized for other systems as an important mechanism by which promiscuity can be achieved.[27] Surprisingly, to our knowledge, no further studies addressed the relation between enzyme dynamics and promiscuity of more than three CYP isoforms comparatively. Furthermore, the approaches did not provide a quantitative relation between substrate scope and structural rigidity of CYP isoforms involved in metabolism. Here, we show for the, to our

Received: May 18, 2021
Published: August 9, 2021

https://doi.org/10.1021/acs.jcim.1c00558
J. Chem. Inf. Model. 2021, 61, 4023−4030

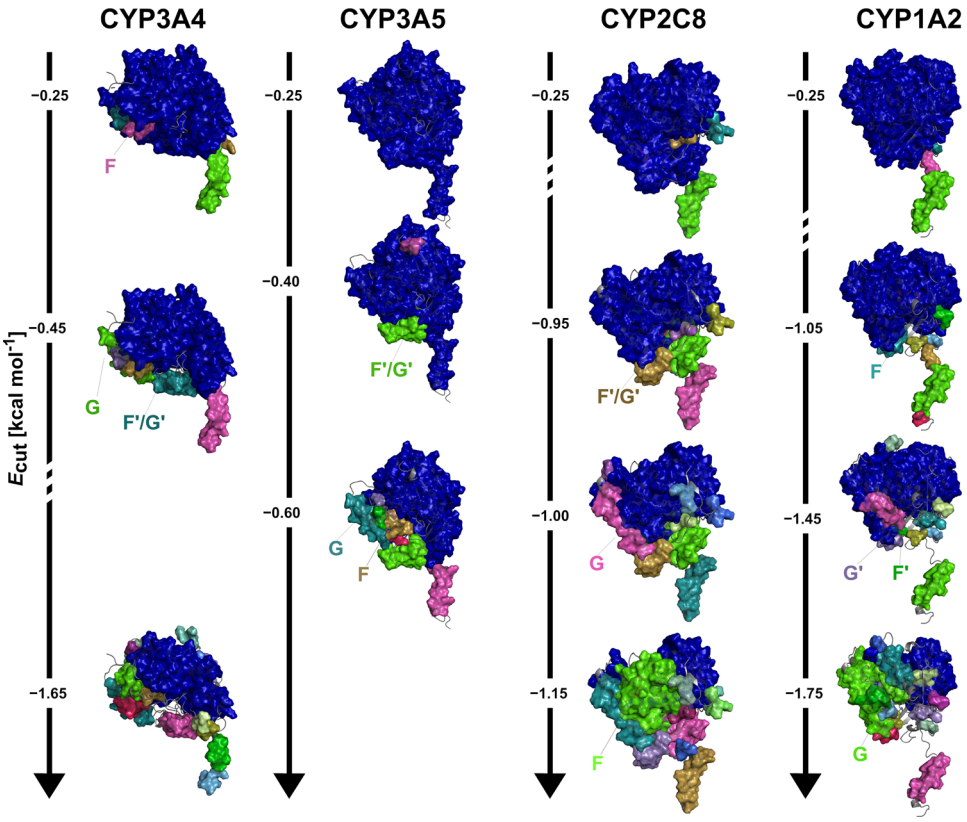


Figure 1. Constraint network analysis of CYP3A4, CYP3A5, CYP2C8, and CYP1A2 and rigid cluster decompositions along constraint dilution trajectories. Rigid clusters of the CYP isoforms at different $\overline{E}_{cut}$ values are colored blue, green, pink, and cyan in descending order of their size. Left: In CYP3A4, the F helix forms its own rigid cluster at $\overline{E}_{cut}$ = −0.25 kcal mol^{-1} and G helix and F′/G′ helix form their own rigid clusters at $\overline{E}_{cut}$ = −0.45 kcal mol^{-1}. At $\overline{E}_{cut}$ = 1.65 kcal mol^{-1}, the largest rigid cluster of CYP3A4 is still covering most of the globular part. Middle-left: In CYP3A5, the F′ helix and G′ helix form their own rigid clusters at $\overline{E}_{cut}$ = −0.40 kcal mol^{-1}, the F helix and the G helix form their own cluster at $\overline{E}_{cut}$ = −0.60 kcal mol^{-1}, and the largest rigid cluster of CYP3A5 is still covering most of the globular part. Middle-right: In CYP2C8, the F′ helix and G′ helix form their own rigid cluster at $\overline{E}_{cut}$ = −0.95 kcal mol^{-1}. The G helix forms its own cluster at $\overline{E}_{cut}$ = −1.00 kcal mol^{-1}. At $\overline{E}_{cut}$ = 1.15 kcal mol^{-1}, the F helix forms its own cluster, and the largest rigid cluster of CYP2C8 is covering only the active site in the globular part. Complete segregation of the F/G region is not recognizable for CYP2C8. Right: In CYP1A2, the F helix forms its own rigid cluster at $\overline{E}_{cut}$ = −1.05 kcal mol^{-1}. The F′ helix and G′ helix form their own cluster at $\overline{E}_{cut}$ = −1.45 kcal mol^{-1}. At $\overline{E}_{cut}$ = 1.75 kcal mol^{-1}, the G helix forms its own cluster, and the largest rigid cluster of CYP1A2 is covering only the active site in the globular part. These examples qualitatively depict that the F/G region of more promiscuous CYP isoforms, such as CYP3A4 and CYP3A5, are less structurally stable than that of more specific isoforms. Furthermore, differences in structural stability are revealed for sequentially close isoforms. The rigid clusters are calculated based on the neighbor stability map of one exemplary MD trajectory, respectively, as described in ref 55.

knowledge, largest data set of CYP isoforms investigated in this context that the structural rigidity of the F/G region is inversely correlated to the enzymes' substrate scope.

MATERIALS AND METHODS

Comparative Modeling of 12 CYP Isoforms. Comparative models of the investigated CYP isoforms were generated with TopModel using the Protect Templates mode.[28] Specified templates that were not removed during threading are listed in Table S1. The globular part and the transmembrane helix were modeled separately, and their positions on or in the membrane were predicted with CCTop web server.[29,30] Afterward, they were docked together using the predicted membrane positions as a spatial restraint. Target sequences were collected from UniProt (Table S1).[31] The quality of the models was assessed with TopScoreSingle (for details, see Text S1).[32] Protonation states of protein residues were adjusted according to pH 7.4 using the Epik routine[33,34] in Maestro.[35] The heme group was transferred from crystal structures that were specified for the TopModel run by aligning the model in Pymol and copying the heme into the model (Table S1).

Conformational Sampling. To improve the robustness of the analyses and quantify the statistical uncertainty of the results, we carried out CNA on ensembles of network topologies generated from five MD trajectories of 1 μs length for each of the enzymes. For this, the generated structural models were embedded by PACKMOL-Memgen[36] into a membrane with a composition of CHL:DOPC:DSPC:DAPC:-DOPE 10:22:13:19:21, the main lipid components of the human endoplasmic reticulum.[37] The GPU particle mesh Ewald implementation[38] of the AMBER18[39] molecular simulations suite was used with ff14SB parameters[40] for the protein, Lipid17 parameters[41] for the membrane, and OPC as a water model.[42] Parameters for the heme and cysteine residues forming the S−Fe bridge between heme and protein were taken from Shahrokh et al.[43] Because covalent bonds to hydrogens were constrained with the SHAKE algorithm,[44] a time step of 2 fs was used. The cutoff for nonbonded interactions was set to 10 Å. For further details, see Text S2.[45] The structures overall (without the transmembrane helix, as it

https://doi.org/10.1021/acs.jcim.1c00558
J. Chem. Inf. Model. 2021, 61, 4023−4030

moves relative to the globular part) and the F/G regions specifically remain stable with backbone RMSD < 4 Å compared to the starting structures (Figures S1−S4).

Constraint Network Analysis. We analyzed static properties, i.e., structural rigidity and its opposite flexibility,[46] of the 12 CYP isoforms predominantly involved in metabolism. The enzymes were represented as constraint networks, where atoms are the nodes and covalent and noncovalent bonds constitute constraints in between.[47] Noncovalent interactions such as hydrogen bonds, salt bridges, hydrophobic tethers, and stacking interactions contribute most to biomolecular stability. The strength of hydrogen bonds and salt bridges was quantified with an empirical energy function.[48] By gradually removing these polar noncovalent constraints from an initial network representation of a biomolecule according to a cutoff energy E_{cut}, a succession of network states σ is generated that forms a "constraint dilution trajectory".[49,50] For this, hydrogen bonds and salt bridges are removed in the order of increasing strength such that for a network state σ, only those hydrogen bonds are kept that have an energy $E_{HB} \leq E_{cut}(\sigma)$. Performing rigidity analysis[51] on such a trajectory reveals a hierarchy of structural stability that reflects the modular structure of biomolecules in terms of secondary, tertiary, and supertertiary structure.

Using a per-residue decomposition scheme to identify the extent to which single residues contribute to the structural stability, we derived neighbor stability maps ($rc_{ij,neighbor}(E_{cut}(\sigma))$) that contain information accumulated over all network states σ along the trajectory[52,53] in that they monitor the persistence of rigid contacts for pairs of residues during a constraint dilution process. In the neighbor stability map, for all residue pairs, the E_{cut} value is given at which the rigid contact between two residues is lost, i.e., when these two residues stop belonging to the same rigid cluster.

The dilution process was applied to each protein conformation of a trajectory separately, and the results were averaged to obtain one neighbor stability map per trajectory. In this work, we consider one trajectory as one independent experiment. A chemical potential energy $E_{i,CNA}$ of residue i is then obtained by summation over all rigid contacts (eq S1). Furthermore, we calculated $\overline{E}_{region,CNA}$ as an average over the $\overline{E}_{i,CNA}$ of all residues of a specific region in the protein or even the entire protein. (eq S2). $\overline{E}_{region,CNA}$ was calculated for each MD trajectory separately and then averaged over all replicas. The computations were done with the Constraint Network Analysis (CNA) program (version 4.0) developed by us,[50] which has been applied in the context of protein thermostability,[54−56] allosteric signaling,[53,57] and substitution influences on the function[58,59] before.

Statistical Analysis and Fitting. Curve fitting was performed with the SciPy module stats.[60] The function linregress calculates linear least-squares regression of two sets of measurement and applies a Wald test whose null hypothesis is that the slope is zero.

RESULTS

Exemplarily, constraint dilution trajectories of CYP3A4, CYP3A5, CYP2C8, and CYP1A2 are shown in Figure 1. Considering that segregation of substructural parts at lower $\overline{E}_{cut}$ values indicates that these parts are more weakly coupled to the remainder of the CYP structure, a qualitative ranking of the CYP isoforms is revealed: In CYP3A4, the F/G region is the first part of the globular domain to segregate at $\overline{E}_{cut}$ = −0.25 to −0.45 kcal mol^{-1} from the largest rigid cluster. In contrast to all other helices of the globular domain, which show higher $\overline{E}_{i,CNA}$ values compared to the TM helix, the F/G region shows $\overline{E}_{i,CNA}$ values comparable to the TM helix, which does not interact with the globular domain (Figure 2).

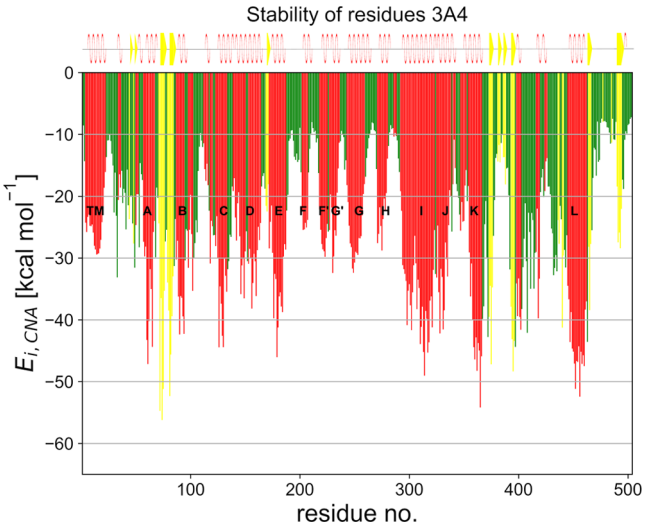


Figure 2. Residue-wise potential chemical energy $\overline{E}_{i,CNA}$ of CYP3A4. $\overline{E}_{i,CNA}$ (eq S1) due to all rigid contacts in which a residue is involved. Secondary structure elements identified with DSSP[80] are colored in red for helices, yellow for β-sheets, and green for loops and given for helices and β-sheets above the plot. Helices are named according to Williams et al.[14]

This is in accordance with findings for different enzymes where helices at the substrate entrances are also less stable, e.g., in aldolases,[61] phosphate synthases,[62] and peptidases,[63] as well as with previous work on CYP enzymes.[20] In CYP3A5, the segregation appears at slightly lower $\overline{E}_{cut}$ values: the F/G region is the first part of the globular domain to segregate at $\overline{E}_{cut}$ = −0.40 to −0.60 kcal mol^{-1} from the largest rigid cluster. For CYP2C8, segregation of the F/G region occurs in the range of −0.95 to −1.15 kcal mol^{-1} and for CYP1A2 at −1.05 to −1.75 kcal mol^{-1}. Hence, the F/G region is structurally the least stable in CYP3A4 and the most in CYP1A2. It is unknown, however, why within the F/G region sometimes the F, G helices and sometimes the F′, G′ helices segregate first. Finally, the rigidity analysis reveals that the anchor helix segregates at the smallest $\overline{E}_{cut}$ in all four cases. This indicates that it is only weakly coupled to the globular part, in line with observations from our MD simulations and ref 64 that the globular part can move relative to the anchor helix (Table S2).

Differences in the structural stability of the F/G region do not only occur between sequentially different CYP isoforms (Figure 1) but also for sequentially close ones as exemplarily depicted for CYP3A4 and CYP3A5 (sequence identity (similarity) 84% (90%) (Figures 1 and 3A)). There, sequence differences particularly involving nonconservative substitutions also occur in the F/G region and lead to higher $\overline{E}_{i,CNA}$ values for the F helix and the G helix of CYP3A5 (Figure 3A,B). As a result, the entire F/G region in CYP3A5 is significantly (p = 0.05, standard independent two-sample t-test) more strongly connected to the remainder of the globular domain than in CYP3A4 (Table S3 and Figure 3C). For the F′/G′ region, no such difference is found, indicating that the small sequence differences there (Figure 3A), including exchanges of

https://doi.org/10.1021/acs.jcim.1c00558
J. Chem. Inf. Model. 2021, 61, 4023−4030

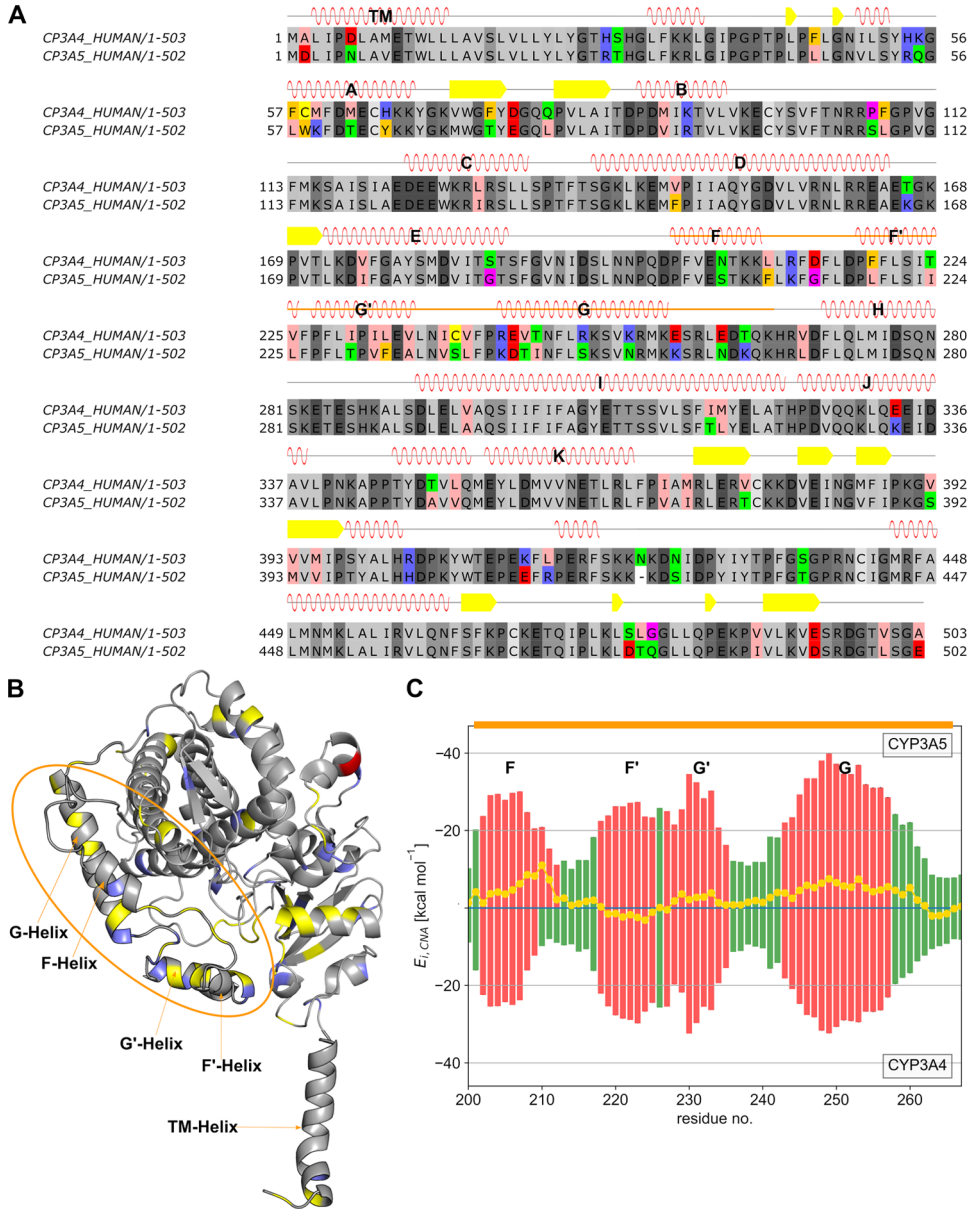


Figure 3. Comparison of CYP3A4 and CYP3A5 with respect to sequence differences. In all panels, the F/G region is marked in orange. (A) Sequence alignment of CYP3A4 and CYP3A5; conserved residues are in gray and nonconserved residues are colored according to the zappo color scheme as to the chemical properties of the side chains. (B) Projection of sequence differences from (A) onto the structure of CYP3A4. Gray: same residue type, blue: same chemical properties, yellow: different chemical properties, and red: gap in CYP3A5. Helices are named according to Williams et al.[14] (C) $\overline{E}_{i,CNA}$ (eq S1) of both isoforms in the F/G region (residue 201−267). More negative values indicate a higher rigidity. The yellow line depicts the difference between both isoforms. Differences in the sequences of the F and G helices cause an overall stronger connection to the remainder of the globular domain of CYP3A5, as indicated by the higher $\overline{E}_{i,CNA}$ values. Differences in the sequences of the F′ and G′ helices do not lead to differences in $\overline{E}_{i,CNA}$. The SEM is <0.05 kcal mol^{-1} in all cases and not depicted. In (A) and (C), secondary structure elements identified with DSSP[80] are colored in red for helices, yellow for β-sheets, and green for loops.

amino acids with those of similar chemical properties, do not lead to differences in the rigidity of this region.

To conclude, in all CYP isoforms exemplarily shown, the F/G region is most weakly coupled to the remainder of the globular domain. Between both sequentially different isoforms and sequentially close ones, differences in the structural stability of this region are revealed by rigidity analysis. These differences qualitatively relate to the substrate scope of the isoforms: more promiscuous isoforms such as CYP3A4 (Table S3) show less structurally stable F/G regions and vice versa.

To quantify the relation between substrate scope and structural rigidity, we computed $\overline{E}_{FG,CNA}$ (eq S2) as a measure for how well residues in the F/G region form rigid contacts (Table S3) and correlated it to the promiscuity index I_{cat} (eq S3) introduced by Nath and Atkins (Table S4).[65] I_{cat} is based on experimentally determined k_{cat} and K_M values for 55 substrates and considers the catalytic efficiency $\overline{e}_i = k_{cat,I}/K_{M,i}$ of an isoform with respect to a substrate i. In this approach, information entropy is used to describe how an enzyme's catalytic efficiency evolved toward different substrates, as measured by the probability p_i that a substrate i will be

https://doi.org/10.1021/acs.jcim.1c00558
J. Chem. Inf. Model. 2021, 61, 4023−4030

metabolized by the enzyme ($p_i = e_i / \sum_{i=1}^{N} e_i$).[65] The index is normalized and ranges from 0 (specific) to 1 (promiscuous).[66]

A good and significant inverse correlation ($R^2 = 0.85$ and $p < 0.01$, Wald test) between the isoform promiscuity and the structural stability of the F/G region is obtained for eight CYP isoforms for which k_{cat} and K_M values are available in ref 66 (Figure 4). Notably, no significant correlations are found if

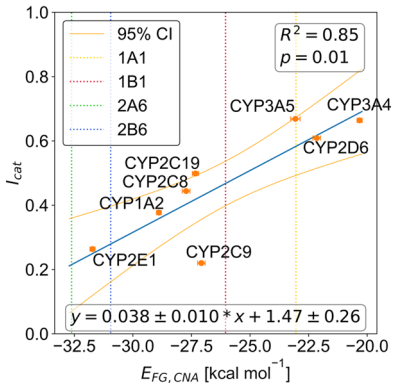


Figure 4. Promiscuity of CYP isoforms is inversely correlated to the structural rigidity of their F/G regions. Correlation between promiscuity index (I_{cat}) and $\overline{E}_{FG,CNA}$ (eq S2) for CYP1A2, CYP2C8, CYP2C9, CYP2C19, CYP2D6, CYP2E1, CYP3A4, and CYP3A5. As to CYP2C9, considered an outlier, the globular part moved away from the membrane during the MD simulations. Vertical lines indicate calculated $\overline{E}_{FG,CNA}$ values for CYP isoforms for which no experimental data for computing I_{cat} are available. The 95% confidence interval limits are shown in orange. Error bars denote the SEM obtained from five, respectively, ten for CYP2E1 and CYP3A4, per-trajectory results. The equation of the predictive line is given at the bottom of the diagram.

$\overline{E}_{region,CNA}$ (eq S2) of the whole enzymes (Figure S7 and Table S5) or of other regions (the I helix, which is the longest helix in the center of CYP enzymes (Figure S8 and Table S6), or the B/C region, which is also part of the main entrance channel (Figure S9 and Table S7)) are used.

Using the correlation for the F/G region, $\overline{E}_{FG,CNA}$ values computed for CYP1A1, CYP1B1, CYP2A6, and CYP2B6 relate to predicted I_{cat} values of 0.58, 0.47, 0.22, and 0.28, respectively, which classifies these CYP isoforms as being lowly to moderately promiscuous. CYP3A5 generally catalyzes the same reactions as CYP3A4 but almost always at lower rates (Figure S12), which leads to similar I_{cat} values that are associated with a slightly lower $\overline{E}_{FG,CNA}$ of CYP3A5 than CYP3A4 (Figure 4). CYP2C9 deviates the most from the correlation line (Figure 4), which may be caused by the globular part that moved away from the membrane during the MD simulations. This is probably due to strain in the loop in the starting structure that connects the globular and transmembrane domains.

DISCUSSION

Our results demonstrate that the structural rigidity of the F/G region quantitatively and inversely correlates to the promiscuity of CYP isoforms involved in metabolism. Previously, enzyme structural dynamics has been recognized as an important mechanism by which promiscuity can be achieved,[27] although the size and architecture of the active site may be further determinants of substrate promiscuity.[67] In addition to CYP enzymes, the possibility of dynamically restructuring active sites has also been recognized for other systems underlying their promiscuity.[68−71] Still, examples for the opposite, i.e., conformational changes selected in evolution such that they enhance specificity in molecular recognition, have also been described in ref 72.

We exploited comprehensive experimental information on the substrate promiscuity of CYP isoforms[66] together with computationally efficient rigidity analyses[46,50,73] of comparative models of the isoforms to understand the molecular origin of the observed promiscuity range. Although I_{cat} used here is a functional parameter defined for a specified set of substrates, promiscuity indices for different enzymes are quantitatively comparable if they have been calculated using the same substrate set.[65] Furthermore, substrates that are chemically similar to each other are expected to be metabolized similarly by a CYP isoform; such correlations in the substrate set would reduce the effective CYP promiscuity. Therefore, we computed the mean maximum pairwise Tanimoto−Combo distance score δ_i of a substrate i to all other substrates in the data set, which ranges from 0 for identical substrates to 2 for dissimilar ones (for details see Text S6); the Tanimoto−Combo distance score accounts for shape and chemical complementary between 3D structures as determined by the Rapid Overlay of Chemical Structures approach. The negatively skewed histogram of δ_i peaks at 1.20, with an average value $\langle\delta\rangle$ of 1.26 (Figure S14), indicating that a substrate is generally more dissimilar than similar to all others in the data set. For comparison, for randomly drawn compounds from the ZINC database, $\langle\delta\rangle$ peaks at 1.12, indicating that such compounds are more similar to each other than those in our data set (Figure S15).

For CYP enzymes, crystallographic studies and other molecular simulations also demonstrated that more promiscuous CYPs show larger structural plasticity and mobility.[74] However, in these studies, only a few isoforms were compared,[15,20,75] no quantitative relations were derived,[15] or short and likely nonconverged MD simulations were applied.[76] Indirectly, the role of structural plasticity and mobility was also investigated in studies focussing on substrate channels.[76] These studies showed qualitative differences in channel properties of CYP isoforms[77−79] but did not take all 12 isoforms involved in metabolism into account and did not derive a quantitative model.

Our quantitative relation allows predicting the promiscuity for CYP1A1, CYP1B1, CYP2A6, and CYP2B6. The difference in $\overline{E}_{FG,CNA}$ between CYP1A1 and CYP1A2 is larger than that observed for other isoforms of the same subfamily such as CYP3A4 and CYP3A5, but still in the range found for CYP2C8 and CYP2C19. Apparently, isoforms of the same subfamily can differ markedly in their substrate scopes, and our approach is able to detect that. However, the model does not allow predicting if a substrate is metabolized by a specific CYP.

In summary, our results signify that characterizing the structural rigidity of the F/G region can be used to classify CYP isoforms involved in metabolism with respect to their substrate scope. Our model may allow predicting the substrate promiscuity of novel CYP enzymes, e.g., found in metagenome approaches for potential use in bioorganic chemistry or biotechnology.

https://doi.org/10.1021/acs.jcim.1c00558
J. Chem. Inf. Model. 2021, 61, 4023−4030

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.1c00558.

Supporting materials and methods, supporting tables, supporting figures, and supporting references (PDF) A repository with used scripts, MD simulation data, and CNA results are available from researchdata.hhu.de via DOI: 10.25838/d5p-21 (TAR.GZ).

■ AUTHOR INFORMATION

Corresponding Author

Holger Gohlke − *Mathematisch-Naturwissenschaftliche Fakultät, Institut für Pharmazeutische und Medizinische Chemie, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany; John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), Institute of Biological Information Processing (IBI-7: Structural Biochemistry), and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany;* orcid.org/0000-0001-8613-1447; Phone: (+49) 211 81 13662; Email: gohlke@uni-duesseldorf.de; Fax: (+49) 211 81 13847

Authors

Daniel Becker − *Mathematisch-Naturwissenschaftliche Fakultät, Institut für Pharmazeutische und Medizinische Chemie, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany*

Prasad V. Bharatam − *Department of Medicinal Chemistry, National Institute of Pharmaceutical Education and Research (NIPER), Mohali 160062 Punjab, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.1c00558

Author Contributions

H.G. designed the study; D.B. performed computations; D.B. and H.G. analyzed results; D.B. and H.G. wrote the manuscript; P.V.B. revised the manuscript; and H.G. and P.V.B. secured funding.

Notes

The authors declare no competing financial interest.

All MD input structures, MD logfiles, CNA results, and scripts used to analyze CNA results are provided in the supporting repository available at researchdata.hhu.de/DOI: 10.25838/d5p-21. Further data sets generated and analyzed during the current study are available from the corresponding author on reasonable request. For molecular simulations, the AMBER18 package of molecular simulation codes was used. AMBER18 is available from http://ambermd.org/. For rigidity analysis, the in-house CNA software package v4.0 was used. CNA is available from the corresponding author free of charge for non-for-profit organizations. Furthermore, a web server is available at https://cpclab.uni-duesseldorf.de/cna/.

■ ACKNOWLEDGMENTS

The authors are grateful for computational support and infrastructure provided by the "Zentrum für Informations- und Medientechnologie" (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user ID: HKF7, CYP450). The study was supported by Bundesministerium für Bildung und Forschung, Germany (BMBF) and the Department of Biotechnology, India (DBT) through project ReMetaDrug (Funding Number 01DQ19002).

■ REFERENCES

(1) Rendic, S.; Guengerich, F. P. Survey of Human Oxidoreductases and Cytochrome P450 Enzymes Involved in the Metabolism of Xenobiotic and Natural Chemicals. *Chem. Res. Toxicol.* **2015**, *28*, 38−42.

(2) Nebert, D. W.; Russell, D. W. Clinical Importance of the Cytochromes P450. *Lancet* **2002**, *360*, 1155−1162.

(3) Ortiz de Montellano, P. R. Hydrocarbon Hydroxylation by Cytochrome P450 Enzymes. *Chem. Rev.* **2010**, *110*, 932−948.

(4) Guengerich, F. P. Common and Uncommon Cytochrome P450 Reactions Related to Metabolism and Chemical Toxicity. *Chem. Res. Toxicol.* **2001**, *14*, 611−650.

(5) Anzenbacher, P.; Anzenbacherová, E. Cytochromes P450 and Metabolism of Xenobiotics. *Cell. Mol. Life Sci.* **2001**, *58*, 737−747.

(6) Black, S. D. Membrane Topology of the Mammalian P450 Cytochromes. *FASEB J.* **1992**, *6*, 680−685.

(7) Cojocaru, V.; Winn, P. J.; Wade, R. C. The Ins and Outs of Cytochrome P450s. *Biochim. Biophys. Acta* **2007**, *1770*, 390−401.

(8) Nelson, D. R. Introductory Remarks on Human Cyps. *Drug Metab. Rev.* **2002**, *34*, 1−5.

(9) Ghosh, D.; Griswold, J.; Erman, M.; Pangborn, W. Structural Basis for Androgen Specificity and Oestrogen Synthesis in Human Aromatase. *Nature* **2009**, *457*, 219−223.

(10) Rieger, M. A.; Ebner, R.; Bell, D. R.; Kiessling, A.; Rohayem, J.; Schmitz, M.; Temme, A.; Rieber, E. P.; Weigle, B. Identification of a Novel Mammary-Restricted Cytochrome P450, Cyp4z1, with Over-expression in Breast Carcinoma. *Cancer Res.* **2004**, *64*, 2357.

(11) Liu, J.; Machalz, D.; Wolber, G.; Sorensen, E. J.; Bureik, M. New Proluciferin Substrates for Human Cyp4 Family Enzymes. *Appl. Biochem. Biotechnol.* **2021**, *193*, 218−237.

(12) Hritz, J.; de Ruiter, A.; Oostenbrink, C. Impact of Plasticity and Flexibility on Docking Results for Cytochrome P450 2d6: A Combined Approach of Molecular Dynamics and Ligand Docking. *J. Med. Chem.* **2008**, *51*, 7469−7477.

(13) Arfeen, M.; Patel, D. S.; Abbat, S.; Taxak, N.; Bharatam, P. V. Importance of Cytochromes in Cyclization Reactions: Quantum Chemical Study on a Model Reaction of Proguanil to Cycloguanil. *J. Comput. Chem.* **2014**, *35*, 2047−2055.

(14) Williams, P. A.; Cosme, J.; Vinković, D. M.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. Crystal Structures of Human Cytochrome P450 3a4 Bound to Metyrapone and Progesterone. *Science* **2004**, *305*, 683.

(15) Hendrychová, T.; Anzenbacherová, E.; Hudeček, J.; Skopalík, J.; Lange, R.; Hildebrandt, P.; Otyepka, M.; Anzenbacher, P. Flexibility of Human Cytochrome P450 Enzymes: Molecular Dynamics and Spectroscopy Reveal Important Function-Related Variations. *Biochim. Biophys. Acta* **2011**, *1814*, 58−68.

(16) Otyepka, M.; Berka, K.; Anzenbacher, P. Is There a Relationship between the Substrate Preferences and Structural Flexibility of Cytochromes P450? *Curr. Drug Metab.* **2012**, *13*, 130−142.

(17) Gora, A.; Brezovsky, J.; Damborsky, J. Gates of Enzymes. *Chem. Rev.* **2013**, *113*, 5871−5923.

(18) Kokkonen, P.; Bednar, D.; Pinto, G.; Prokop, Z.; Damborsky, J. Engineering Enzyme Access Tunnels. *Biotechnol. Adv.* **2019**, *37*, No. 107386.

(19) Kingsley, L. J.; Lill, M. A. Substrate Tunnels in Enzymes: Structure−Function Relationships and Computational Methodology. *Proteins* **2015**, *83*, 599−611.

(20) Skopalík, J.; Anzenbacher, P.; Otyepka, M. Flexibility of Human Cytochromes P450: Molecular Dynamics Reveals Differences

between Cyps 3a4, 2c9, and 2a6, Which Correlate with Their Substrate Preferences. *J. Phys. Chem. B* **2008**, *112*, 8165−8173.
(21) Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D. Enzyme Dynamics During Catalysis. *Science* **2002**, *295*, 1520.
(22) Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D. A Hierarchy of Timescales in Protein Dynamics Is Linked to Enzyme Catalysis. *Nature* **2007**, *450*, 913−916.
(23) Holliday, M. J.; Camilloni, C.; Armstrong, G. S.; Vendruscolo, M.; Eisenmesser, E. Z. Networks of Dynamic Allostery Regulate Enzyme Function. *Structure* **2017**, *25*, 276−286.
(24) Goodey, N. M.; Benkovic, S. J. Allosteric Regulation and Catalysis Emerge Via a Common Route. *Nat. Chem. Biol.* **2008**, *4*, 474−482.
(25) Saavedra, H. G.; Wrabl, J. O.; Anderson, J. A.; Li, J.; Hilser, V. J. Dynamic Allostery Can Drive Cold Adaptation in Enzymes. *Nature* **2018**, *558*, 324−328.
(26) Tzeng, S.-R.; Kalodimos, C. G. Protein Dynamics and Allostery: An Nmr View. *Curr. Opin. Struct. Biol.* **2011**, *21*, 62−67.
(27) Nobeli, I.; Favia, A. D.; Thornton, J. M. Protein Promiscuity and Its Implications for Biotechnology. *Nat. Biotechnol.* **2009**, *27*, 157−167.
(28) Mulnaes, D.; Porta, N.; Clemens, R.; Apanasenko, I.; Reiners, J.; Gremer, L.; Neudecker, P.; Smits, S. H. J.; Gohlke, H. Topmodel: Template-Based Protein Structure Prediction at Low Sequence Identity Using Top-Down Consensus and Deep Neural Networks. *J. Chem. Theory Comput.* **2020**, *16*, 1953−1967.
(29) Dobson, L.; Reményi, I.; Tusnády, G. E. Cctop: A Consensus Constrained Topology Prediction Web Server. *Nucleic Acids Res.* **2015**, *43*, W408−W412.
(30) Dobson, L.; Reményi, I.; Tusnády, G. E. The Human Transmembrane Proteome. *Biol. Direct* **2015**, *10*, No. 31.
(31) UniProt, C. Uniprot: A Worldwide Hub of Protein Knowledge. *Nucleic Acids Res.* **2019**, *47*, D506−D515.
(32) Mulnaes, D.; Gohlke, H. Topscore: Using Deep Neural Networks and Large Diverse Data Sets for Accurate Protein Model Quality Assessment. *J. Chem. Theory Comput.* **2018**, *14*, 6117−6126.
(33) Greenwood, J. R.; Calkins, D.; Sullivan, A. P.; Shelley, J. C. Towards the Comprehensive, Rapid, and Accurate Prediction of the Favorable Tautomeric States of Drug-Like Molecules in Aqueous Solution. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 591−604.
(34) Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: A Software Program for Pkaprediction and Protonation State Generation for Drug-Like Molecules. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 681−691.
(35) *Schrödinger Release 2019-3: Maestro*; Schrödinger, LLC: New York, NY, 2019.
(36) Schott-Verdugo, S.; Gohlke, H. Packmol-Memgen: A Simple-to-Use, Generalized Workflow for Membrane-Protein-Lipid-Bilayer System Building. *J. Chem. Inf. Model.* **2019**, *59*, 2522−2528.
(37) Yeagle, P. *The Membranes of Cells*, 3rd ed.; Elsevier: Amsterdam, 2016; Vol. *xii*, 439 pp.
(38) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.
(39) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, I. T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Giambasu, G.; Giese, T.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Krasny, R.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Man, V.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wilson, L.; Wolf, R. M.; Wu, X.; Xiao, L.; Xiong, Y.; York, D. M.; Kollman, P. A. *Amber 2019*; University of California: San Francisco, 2019.
(40) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14sb: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99sb. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(41) Dickson, C. J.; Madej, B. D.; Skjevik, A. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.* **2014**, *10*, 865−879.
(42) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863−3871.
(43) Shahrokh, K.; Orendt, A.; Yost, G. S.; Cheatham, T. E., 3rd Quantum Mechanically Derived Amber-Compatible Heme Parameters for Various States of the Cytochrome P450 Catalytic Cycle. *J. Comput. Chem.* **2012**, *33*, 119−133.
(44) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(45) Rathi, P. C.; Radestock, S.; Gohlke, H. Thermostabilizing Mutations Preferentially Occur at Structural Weak Spots with a High Mutation Ratio. *J. Biotechnol.* **2012**, *159*, 135−144.
(46) Hermans, S. M. A.; Pfleger, C.; Nutschel, C.; Hanke, C. A.; Gohlke, H. Rigidity Theory for Biomolecules: Concepts, Software, and Applications. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2017**, *7*, No. e1311.
(47) Jacobs, D. J.; Rader, A. J.; Kuhn, L. A.; Thorpe, M. F. Protein Flexibility Predictions Using Graph Theory. *Proteins* **2001**, *44*, 150−165.
(48) Dahiyat, B. I.; Benjamin Gordon, D.; Mayo, S. L. Automated Design of the Surface Positions of Protein Helices. *Protein Sci.* **1997**, *6*, 1333−1337.
(49) Radestock, S.; Gohlke, H. Exploiting the Link between Protein Rigidity and Thermostability for Data-Driven Protein Engineering. *Eng. Life Sci.* **2008**, *8*, 507−522.
(50) Pfleger, C.; Rathi, P. C.; Klein, D. L.; Radestock, S.; Gohlke, H. Constraint Network Analysis (Cna): A Python Software Package for Efficiently Linking Biomacromolecular Structure, Flexibility, (Thermo-)Stability, and Function. *J. Chem. Inf. Model.* **2013**, *53*, 1007−1015.
(51) Jacobs, D. J.; Thorpe, M. F. Generic Rigidity Percolation: The Pebble Game. *Phys. Rev. Lett.* **1995**, *75*, 4051−4054.
(52) Pfleger, C.; Radestock, S.; Schmidt, E.; Gohlke, H. Global and Local Indices for Characterizing Biomolecular Flexibility and Rigidity. *J. Comput. Chem.* **2013**, *34*, 220−233.
(53) Pfleger, C.; Minges, A.; Boehm, M.; McClendon, C. L.; Torella, R.; Gohlke, H. Ensemble- and Rigidity Theory-Based Perturbation Approach to Analyze Dynamic Allostery. *J. Chem. Theory Comput.* **2017**, *13*, 6343−6357.
(54) Contreras, F.; Nutschel, C.; Beust, L.; Davari, M. D.; Gohlke, H.; Schwaneberg, U. Can Constraint Network Analysis Guide the Identification Phase of Knowvolution? A Case Study on Improved Thermostability of an Endo-B-Glucanase. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 743−751.
(55) Rathi, P. C.; Fulton, A.; Jaeger, K.-E.; Gohlke, H. Application of Rigidity Theory to the Thermostabilization of Lipase a from *Bacillus subtilis*. *PLoS Comput. Biol.* **2016**, *12*, No. e1004754.
(56) Nutschel, C.; Fulton, A.; Zimmermann, O.; Schwaneberg, U.; Jaeger, K.-E.; Gohlke, H. Systematically Scrutinizing the Impact of Substitution Sites on Thermostability and Detergent Tolerance for Bacillus Subtilis Lipase A. *J. Chem. Inf. Model.* **2020**, *60*, 1568−1584.
(57) Pfleger, C.; Kusch, J.; Kondapuram, M.; Schwabe, T.; Sattler, C.; Benndorf, K.; Gohlke, H. Allosteric Signaling in C-Linker and Cyclic Nucleotide-Binding Domain of Hcn2 Channels. *Biophys. J.* **2021**, *120*, 950−963.
(58) Preising, M. N.; Görg, B.; Friedburg, C.; Qvartskhava, N.; Budde, B. S.; Bonus, M.; Toliat, M. R.; Pfleger, C.; Altmüller, J.; Herebian, D.; Beyer, M.; Zöllner, H. J.; Wittsack, H.-J.; Schaper, J.; Klee, D.; Zechner, U.; Nürnberg, P.; Schipper, J.; Schnitzler, A.; Gohlke, H.; Lorenz, B.; Häussinger, D.; Bolz, H. J. Biallelic Mutation of Human Slc6a6 Encoding the Taurine Transporter Taut Is Linked to Early Retinal Degeneration. *FASEB J.* **2019**, *33*, 11507−11527.

https://doi.org/10.1021/acs.jcim.1c00558
J. Chem. Inf. Model. 2021, 61, 4023−4030

(59) Wifling, D.; Pfleger, C.; Kaindl, J.; Ibrahim, P.; Kling, R. C.; Buschauer, A.; Gohlke, H.; Clark, T. Basal Histamine H4 Receptor Activation: Agonist Mimicry by the Diphenylalanine Motif. *Chem. – Eur. J.* **2019**, *25*, 14613−14624.
(60) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; Vijaykumar, A.; Bardelli, A. P.; Rothberg, A.; Hilboll, A.; Kloeckner, A.; Scopatz, A.; Lee, A.; Rokem, A.; Woods, C. N.; Fulton, C.; Masson, C.; Häggström, C.; Fitzgerald, C.; Nicholson, D. A.; Hagen, D. R.; Pasechnik, D. V.; Olivetti, E.; Martin, E.; Wieser, E.; Silva, F.; Lenders, F.; Wilhelm, F.; Young, G.; Price, G. A.; Ingold, G.-L.; Allen, G. E.; Lee, G. R.; Audren, H.; Probst, I.; Dietrich, J. P.; Silterra, J.; Webber, J. T.; Slavič, J.; Nothman, J.; Buchner, J.; Kulick, J.; Schönberger, J. L.; de Miranda Cardoso, J. V.; Reimer, J.; Harrington, J.; Rodríguez, J. L. C.; Nunez-Iglesias, J.; Kuczynski, J.; Tritz, K.; Thoma, M.; Newville, M.; Kümmerer, M.; Bolingbroke, M.; Tartre, M.; Pak, M.; Smith, N. J.; Nowaczyk, N.; Shebanov, N.; Pavlyk, O.; Brodtkorb, P. A.; Lee, P.; McGibbon, R. T.; Feldbauer, R.; Lewis, S.; Tygier, S.; Sievert, S.; Vigna, S.; Peterson, S.; More, S.; Pudlik, T.; Oshima, T.; Pingel, T. J.; Robitaille, T. P.; Spura, T.; Jones, T. R.; Cera, T.; Leslie, T.; Zito, T.; Krauss, T.; Upadhyay, U.; Halchenko, Y. O.; Vázquez-Baeza, Y.; SciPy, C.; et al. Scipy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261−272.
(61) Dick, M.; Weiergraber, O. H.; Classen, T.; Bisterfeld, C.; Bramski, J.; Gohlke, H.; Pietruszka, J. Trading Off Stability against Activity in Extremophilic Aldolases. *Sci. Rep.* **2016**, *6*, No. 17908.
(62) Merz, A.; Yee, M.-c.; Szadkowski, H.; Pappenberger, G.; Crameri, A.; Stemmer, W. P. C.; Yanofsky, C.; Kirschner, K. Improving the Catalytic Activity of a Thermophilic Enzyme at Low Temperatures. *Biochemistry* **2000**, *39*, 880−889.
(63) Fuxreiter, M.; Magyar, C.; Juhász, T.; Szeltner, Z.; Polgár, L.; Simon, I. Flexibility of Prolyl Oligopeptidase: Molecular Dynamics and Molecular Framework Analysis of the Potential Substrate Pathways. *Proteins* **2005**, *60*, 504−512.
(64) Berka, K.; Hendrychová, T.; Anzenbacher, P.; Otyepka, M. Membrane Position of Ibuprofen Agrees with Suggested Access Path Entrance to Cytochrome P450 2c9 Active Site. *J. Phys. Chem. A* **2011**, *115*, 11248−11255.
(65) Nath, A.; Atkins, W. M. A Quantitative Index of Substrate Promiscuity. *Biochemistry* **2008**, *47*, 157−166.
(66) Foti, R. S.; Honaker, M.; Nath, A.; Pearson, J. T.; Buttrick, B.; Isoherranen, N.; Atkins, W. M. Catalytic Versus Inhibitory Promiscuity in Cytochrome P450s: Implications for Evolution of New Function. *Biochemistry* **2011**, *50*, 2387−2393.
(67) Müller, H.; Becker, A. K.; Palm, G. J.; Berndt, L.; Badenhorst, C. P. S.; Godehard, S. P.; Reisky, L.; Lammers, M.; Bornscheuer, U. T. Sequence-Based Prediction of Promiscuous Acyltransferase Activity in Hydrolases. *Angew. Chem., Int. Ed.* **2020**, *59*, 11607−11612.
(68) Seibert, C. M.; Raushel, F. M. Structural and Catalytic Diversity within the Amidohydrolase Superfamily. *Biochemistry* **2005**, *44*, 6383−6391.
(69) Oppermann, U.; Filling, C.; Hult, M.; Shafqat, N.; Wu, X.; Lindh, M.; Shafqat, J.; Nordling, E.; Kallberg, Y.; Persson, B.; Jörnvall, H. Short-Chain Dehydrogenases/Reductases (SDR): The 2002 Update. *Chem.−Biol. Interact.* **2003**, *143−144*, 247−253.
(70) Fushinobu, S.; Nishimasu, H.; Hattori, D.; Song, H.-J.; Wakagi, T. Structural Basis for the Bifunctionality of Fructose-1,6-Bisphosphate Aldolase/Phosphatase. *Nature* **2011**, *478*, 538−541.
(71) Du, J.; Say, R. F.; Lü, W.; Fuchs, G.; Einsle, O. Active-Site Remodelling in the Bifunctional Fructose-1,6-Bisphosphate Aldolase/Phosphatase. *Nature* **2011**, *478*, 534−537.
(72) Nutschel, C.; Coscolín, C.; David, B.; Mulnaes, D.; Ferrer, M.; Jaeger, K.-E.; Gohlke, H. Promiscuous Esterases Counterintuitively Are Less Flexible Than Specific Ones. *J. Chem. Inf. Model.* **2021**, *61*, 2383−2395.
(73) Krüger, D. M.; Rathi, P. C.; Pfleger, C.; Gohlke, H. CNA Web Server: Rigidity Theory-Based Thermal Unfolding Simulations of Proteins for Linking Structure, (Thermo-)Stability, and Function. *Nucleic Acids Res.* **2013**, *41*, W340−W348.
(74) Hawkins, P. C.; Skillman, A. G.; Nicholls, A. Comparison of Shape-Matching and Docking as Virtual Screening Tools. *J. Med. Chem.* **2007**, *50*, 74−82.
(75) Ekroos, M.; Sjögren, T. Structural Basis for Ligand Promiscuity in Cytochrome P450 3a4. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13682−13687.
(76) Schleinkofer, K.; Sudarko; Winn, P. J.; Lüdemann, S. K.; Wade, R. C. Do Mammalian Cytochrome P450s Show Multiple Ligand Access Pathways and Ligand Channelling? *EMBO Rep.* **2005**, *6*, 584−589.
(77) Kingsley, L. J.; Lill, M. A. Ensemble Generation and the Influence of Protein Flexibility on Geometric Tunnel Prediction in Cytochrome P450 Enzymes. *PLoS One* **2014**, *9*, No. e99408.
(78) Berka, K.; Paloncýová, M.; Anzenbacher, P.; Otyepka, M. Behavior of Human Cytochromes P450 on Lipid Membranes. *J. Phys. Chem. B* **2013**, *117*, 11556−11564.
(79) Urban, P.; Lautier, T.; Pompon, D.; Truan, G. Ligand Access Channels in Cytochrome P450 Enzymes: A Review. *Int. J. Mol. Sci.* **2018**, *19*, No. 1617.
(80) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22*, 2577−2637.

https://doi.org/10.1021/acs.jcim.1c00558
J. Chem. Inf. Model. 2021, 61, 4023−4030