

RESEARCH ARTICLE

# Characterization of an Additional Splice Acceptor Site Introduced into *CYP4B1* in *Hominoidae* during Evolution

Eva M. Schmidt<sup>1</sup>✉, Constanze Wiek<sup>2</sup>✉, Oliver T. Parkinson<sup>3</sup>, Katharina Roellecke<sup>2</sup>, Marcel Freund<sup>2</sup>, Michael Gombert<sup>1</sup>, Nadine Lottmann<sup>2</sup>, Charles A. Steward<sup>4</sup>, Christof M. Kramm<sup>5</sup>, Vladimir Yarov-Yarovoy<sup>6</sup>, Allan E. Rettie<sup>3</sup>, Helmut Hanenberg<sup>2,7,8\*</sup>

**1** Department of Pediatric Hematology, Oncology and Clinical Immunology, Children's Hospital, Heinrich Heine University, 40225 Düsseldorf, Germany, **2** Department of Otorhinolaryngology and Head/Neck Surgery, Heinrich Heine University, 40225 Düsseldorf, Germany, **3** Department of Medicinal Chemistry, School of Pharmacy, University of Washington, Seattle, WA 98195, United States of America, **4** Wellcome Trust Sanger Institute, Cambridge, United Kingdom, **5** Division of Pediatric Hematology and Oncology, Department of Child and Adolescent Health, University of Göttingen, 37099 Göttingen, Germany, **6** Departments of Physiology and Membrane Biology, University of California Davis, Davis, CA 95616, United States of America, **7** Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN 46202, United States of America, **8** Department of Pediatrics III, University Children's Hospital Essen, University of Duisburg-Essen, 45122 Essen, Germany



**OPEN ACCESS**

**Citation:** Schmidt EM, Wiek C, Parkinson OT, Roellecke K, Freund M, Gombert M, et al. (2015) Characterization of an Additional Splice Acceptor Site Introduced into *CYP4B1* in *Hominoidae* during Evolution. PLoS ONE 10(9): e0137110. doi:10.1371/journal.pone.0137110

**Editor:** Xinghui Qiu, Institute of Zoology, Chinese Academy of Sciences, CHINA

**Received:** May 21, 2015

**Accepted:** August 12, 2015

**Published:** September 10, 2015

**Copyright:** © 2015 Schmidt et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** This work was initially supported by the Deutsche José-Carreras-Leukämie-Stiftung e. V. (Munich, Germany), the Deutsche Krebshilfe e. V. (to CMK and HH) and the NIH grant R01 GM49054 (AER). Later support was from the Forschungskommission of the Medical Faculty and the Strategische Forschungskommission of the Heinrich Heine University, Düsseldorf, Germany (to CW), the UW School of Pharmacy Brady Fund for Natural Products Research (to AER), and NIH R01

✉ These authors contributed equally to this work.

\* [Helmut.Hanenberg@uni-due.de](mailto:Helmut.Hanenberg@uni-due.de)

## Abstract

CYP4B1 belongs to the cytochrome P450 family 4, one of the oldest P450 families whose members have been highly conserved throughout evolution. The CYP4 monooxygenases typically oxidize fatty acids to both inactive and active lipid mediators, although the endogenous ligand(s) is largely unknown. During evolution, at the transition of great apes to humanoids, the CYP4B1 protein acquired a serine instead of a proline at the canonical position 427 in the meander region. Although this alteration impairs P450 function related to the processing of naturally occurring lung toxins, a study in transgenic mice suggested that an additional serine insertion at position 207 in human CYP4B1 can rescue the enzyme stability and activity. Here, we report that the genomic insertion of a CAG triplet at the intron 5–exon 6 boundary in human *CYP4B1* introduced an additional splice acceptor site in frame. During evolution, this change occurred presumably at the stage of *Hominoidae* and leads to two major isoforms of the CYP4B1 enzymes of humans and great apes, either with or without a serine 207 insertion (insSer207). We further demonstrated that the CYP4B1 enzyme with insSer207 is the dominant isoform (76%) in humans. Importantly, this amino acid insertion did not affect the 4-ipomeanol metabolizing activities or stabilities of the native rabbit or human CYP4B1 enzymes, when introduced as transgenes in human primary cells and cell lines. In our 3D modeling, this functional neutrality of insSer207 is compatible with its predicted location on the exterior surface of CYP4B1 in a flexible side chain. Therefore, the Ser207 insertion does not rescue the P450 functional activity of human CYP4B1 that has been lost during evolution.

CA155294 (to HH). HH is supported by the Lilly Foundation Physician/Scientist initiative. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** HH may receive royalties based on a license agreement between Indiana University, Indianapolis, USA, and Takara Shuzo Inc., Kyoto, Japan, on the sales and usage of Retronectin CH-296 in transduction protocols. The other authors declare no competing financial interests. This does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

## Introduction

Cytochrome P450s (CYPs) constitute a large superfamily of genes that have co-evolved with their hosts—higher plants, prokaryotic and eukaryotic organisms—from a single common ancestor [1, 2]. In humans, the HUGO Gene Nomenclature Committee currently lists 107 genes and pseudogenes (<http://www.genenames.org/genefamilies/CYP>) belonging to 4 major and 13 smaller P450 families [3]. Although there are remarkably similar substrate specificities for the vast majority of P450 orthologs in mammals due to semi-conserved active-site sequences and ligand access channels [4], human *CYP4B1* stands out as a functionally unusual member of this single-gene subfamily (reviewed in [5]).

*CYP4B1* was initially recognized as the P450 enzyme responsible for the activation of a naturally occurring pro-toxin, 4-ipomeanol (4-IPO), to an alkylating agent [6–10]. 4-IPO is produced by sweet potatoes (*Ipomoea batatas*) infected with the fungus *Fusarium solani* [11–14]. When livestock ingest moldy sweet potatoes, 4-IPO uptake induces selective cellular toxicity in the lungs as the primary organ affected [11, 12, 14] where 70% of all *CYP4B1* transcripts are expressed [15–18]. Based on the preclinical observations in rodents and dogs that doses can be established where the cytotoxicity is restricted to the lungs and the findings that human lung cancer cells *in vitro* and *in vivo* can be specifically targeted (reviewed in [19]), three human clinical phase I/II trials were conducted [20–22]. Surprisingly, dose-escalation studies revealed that no lung toxicity, but some reversible dose-dependent liver toxicity occurred after intravenous administration of 4-IPO. However, no objective anti-tumor effects were noted, for either lung or liver cancer patients [20–22].

Major interspecies differences in the activity of *CYP4B1* enzyme between the human and other mammalian homologs have been reported (reviewed in [5]). When overexpressed in human HepG2 liver cells, the native human *CYP4B1* enzyme is unable to metabolize the classical substrates such as 2-aminoanthracene (2-AA) [23, 24] and 4-IPO [9, 25]. In addition, no catalytic activity of the human enzyme was detected upon its expression in insect cells using a baculovirus system [26], presumably because the human enzyme sequence carries serine instead of proline at position 427 in the meander region near the heme-binding site [26, 27]. Remarkably, the presence of serine instead of proline at position 427 is specific to human *CYP4B1* and also unique among other human *CYP450* enzymes [5, 28].

In addition to the initially identified human cDNA sequence [25], three groups of investigators reported an in-frame insertion in the major *CYP4B1* transcript isolated from several human tissues [29–31]. Bylund *et al.* performed RT-PCR studies on four human seminal vesicle samples and detected a mixture of the native *CYP4B1* and an isoform with a 3-bp insertion (CAG) encoding serine at position 207 in all samples [29]. Carr *et al.* suggested that the serine insertion is the result of an alternatively splice transcript but was not able to find a transcript without serine at position 207 in five bacterial clones amplified from human lung [31]. In these studies, the functional significance of the serine insertion was not determined, because it was not possible to express active human *CYP4B1* enzyme, either with or without the serine insertion [31]. In 2001, Imaoka *et al.* reported that the serine insertion was carried by the *CYP4B1* 'gene' in 50 different individuals [30], suggesting that it is the major transcript in humans, although no supporting data were presented. This group was also able to express human *CYP4B1* insSer207 in the liver of a transgenic mouse and as a fusion protein of *CYP4B1* and NADPH P450 reductase in yeast cells [30]. The microsomal preparations containing the human protein catalyzed  $\omega$ -hydroxylation of lauric acid and activated 2-AF, both known substrates for *CYP4B1*, thus prompting the authors to hypothesize that the serine insertion "may also stabilize the native *CYP4B1*" [30]. However, these findings regarding a functional activity of native human *CYP4B1* with the serine insertion towards classical substrates of animals

orthologs have not been confirmed by others (reviewed in [5]) and are in sharp contrast to the results of phase I/II dose escalation trials in humans in which 4-IPO was administered as an anti-lung cancer agent [20–22]. Although humans express 70% of all *CYP4B1* transcripts in the lung [29–32], no lung toxicity was observed with administration of 4-IPO doses that are highly toxic/lethal in animals [17, 19, 33, 34].

To clarify some of these issues, we recently re-engineered human *CYP4B1* for efficient activation of 4-IPO by introducing amino acid residues from the highly active rabbit enzyme into the native human protein as well as the p.S427P *CYP4B1* mutant [28]. Using lentiviral vector systems to achieve stable expression of human and rabbit *CYP4B1* in human HepG2 liver cells and primary T cells, we demonstrated that native human *CYP4B1* is an inactive and relatively unstable protein with a short half-life compared to its rabbit homolog. Exchange of serine with proline at position 427 improved the half-life of the human *CYP4B1* protein and dramatically increased its enzymatic activity towards 4-IPO, although this single point mutant was still less active than the rabbit enzyme. Eventually, through systematic mutagenesis, we identified 12 additional amino acid substitutions that conferred all of the functional activity of rabbit *CYP4B1* to human *CYP4B1* [28]. These prior efforts establish a baseline against which to probe the functional consequences of insSer207 in *CYP4B1*.

In the present study, we analyzed the molecular basis for serine insertion at position 207 in the human *CYP4B1* enzyme and determined the frequency of insSer207 in human *CYP4B1* transcripts. In order to determine whether insSer207 could have emerged as a compensatory mechanism for reduced P450 enzymatic activity due to a p.P427S exchange that had occurred in humans, we tested whether insSer207 stabilizes the half-life and/or enhances the functional activity of native human *CYP4B1* and the p.S427P *CYP4B1* mutant proteins. Finally, we explored the evolution of the human *CYP4B1* genomic sequence through comparison of insSer207 and p.P427S in humans to those in great apes and non-human primates as well as other mammals.

## Materials and Methods

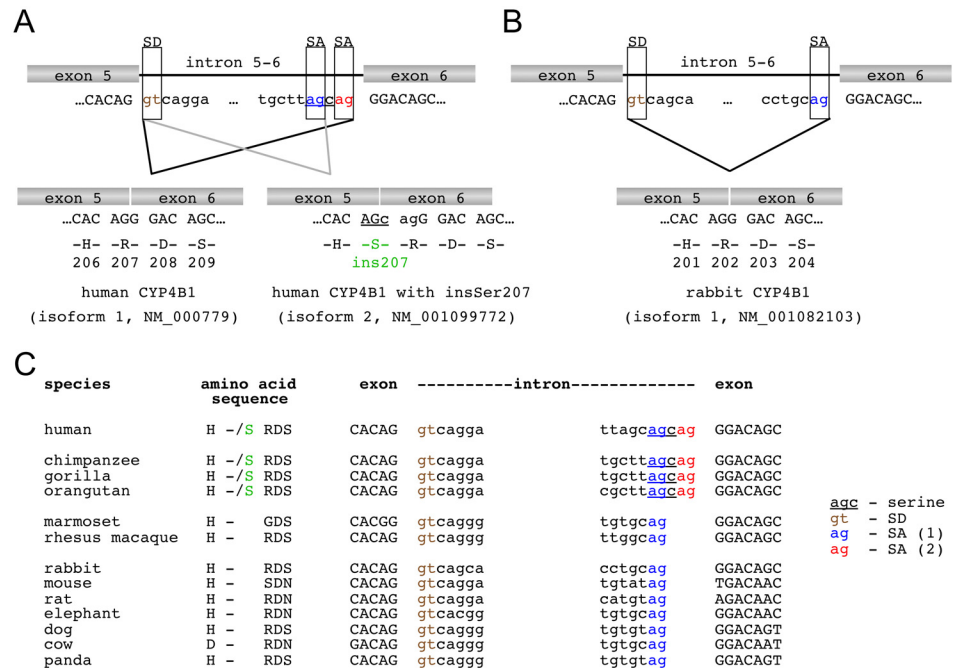
### Serine insertion in normal human *CYP4B1* mRNA

To assess the frequency of the serine insertion in human *CYP4B1* mRNA (Fig 1A), a PCR was performed on human reference cDNA (Clontech #636690, Lot #1005341A) using the exon 4 forward primer 5' –GCC**CTCGAG**GGTTGCAGCACCGCAAGCTG with the *XhoI* restriction enzyme recognition site (bold) and the exon 9 reverse primer 5' –GGG**GCTAGC**CACAGGTGGG TAGAGGCGG with the *NheI* site (Fig 2A). PCR products of 740 and 743 bp were ligated into the multiple cloning site (mcs) of the lentiviral vector IRES-EGFP with an internal ribosome entry site—enhanced green fluorescent protein (Fig 3A) and then transformed into Top10 bacteria (Invitrogen, Karlsruhe, Germany). Plasmid DNA was isolated from 220 bacterial clones and sequenced using the forward primer 5' –GGACCTGAAATGACCCTGCG that binds in the spleen focus-forming virus (SFFV) promoter.

### Cell cultures and primary cells

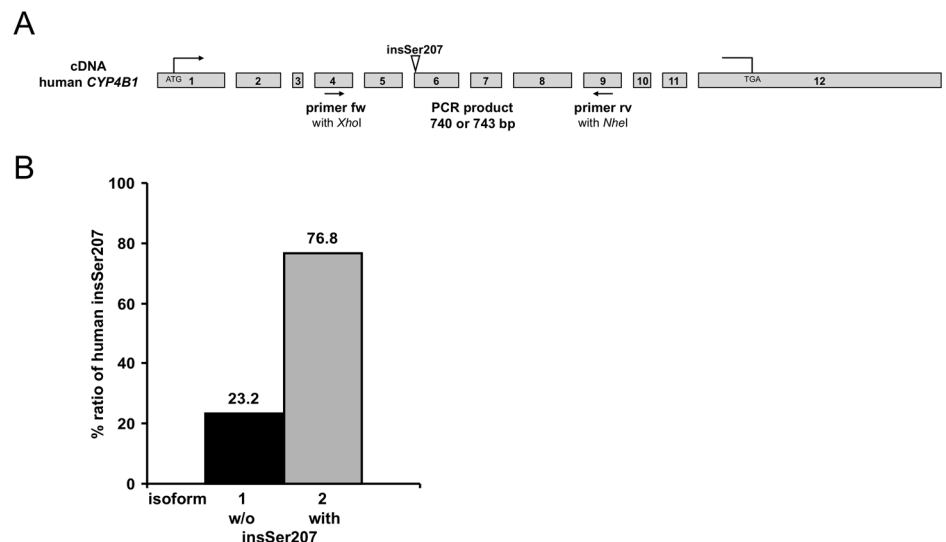
Human embryonic kidney (HEK293T) cells and human hepatoma (HepG2) cells were purchased from ATCC (Manassas, VA), and human fibrosarcoma HT1080 cells were obtained from DSMZ (Braunschweig, Germany). All cells were grown in Dulbecco's Modified Eagle's Medium supplemented with penicillin (100 U/mL), streptomycin (100 µg/mL), 2 mM glutamine (all from Gibco/BRL, Karlsruhe, Germany), and 10% heat-inactivated fetal bovine serum (PAN, Aidenbach, Germany) at 37°C in a humidified atmosphere with 5% CO<sub>2</sub>.

Primary human T lymphocytes from peripheral blood (PB) samples from healthy adult volunteers who provided written informed consent that is documented in the department of



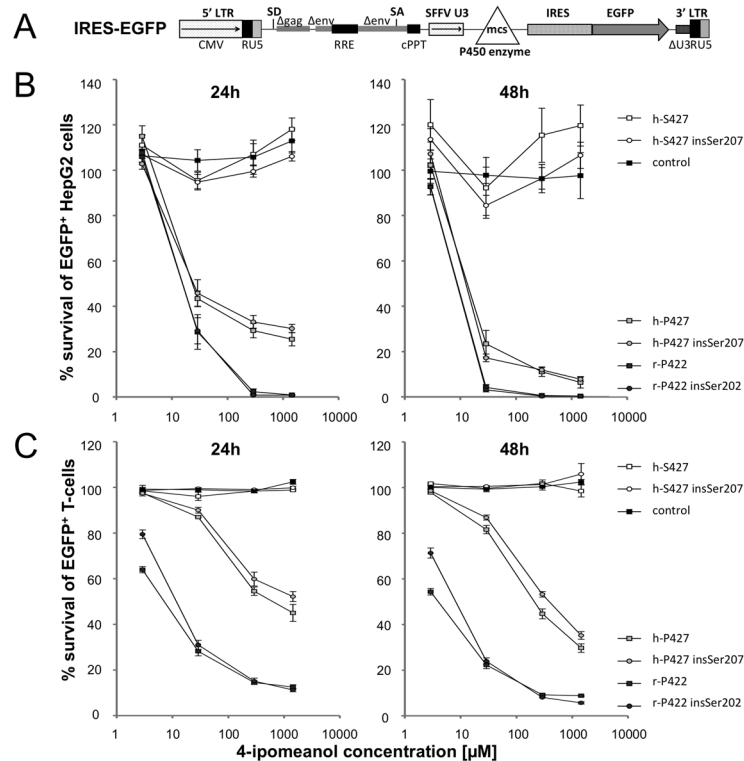
**Fig 1. Alternative Splicing of *CYP4B1*.** (A) Human *CYP4B1*. The human *CYP4B1* intron 5–6 harbors two splice acceptors (SAs) in frame allowing the generation of two alternative *CYP4B1* transcripts, with (isoform 2: NM\_001099772) or without (isoform 1: NM\_000779) insSer207. (B) Rabbit *CYP4B1*. Intron 5–6 of rabbit *CYP4B1* carries only a single SA and therefore generates a single transcript without insSer202 (isoform 1: NM\_001082103). (C) Alignment of *CYP4B1* sequences from other species. Alignment of the genomic DNA of *CYP4B1* from other species shows that the insertion of the CAG triplet, which generates the additional splice acceptor site, exists only in humans and great apes, and not in rhesus, macaque, marmoset or other mammals.

doi:10.1371/journal.pone.0137110.g001



**Fig 2. Frequency of insSer207 in human *CYP4B1* transcripts.** (A) Schematic outline of *CYP4B1*. The forward (fw) primer binds in exon 4 and the reverse (rv) primer in exon 9, thus resulting in PCR products of 240 bp and 243 bp, respectively. (B) Frequency of human *CYP4B1* without (w/o) and with insSer207. The plasmid DNA from 220 bacterial colonies was isolated and sequenced. The percentages of *CYP4B1* transcripts with and without (w/o) insSer207 are shown.

doi:10.1371/journal.pone.0137110.g002



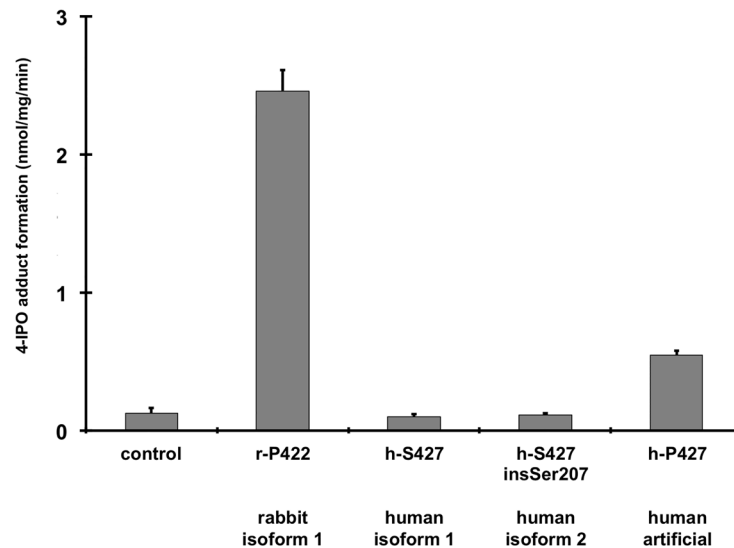
**Fig 3. Functional analysis of CYP4B1.** (A) Schematic outline of the lentiviral (LV) vectors used for expression of CYP4B1. For functional testing, the different CYP4B1 forms were cloned into a LV vector with an IRES-EGFP site. CMV: CMV promoter; SD: splice donor; LTR: long terminal repeat; SA: splice acceptor; RRE: Rev responsive element, cPPT: central polypurine binding tract; SFFV U3: U3 promoter of the spleen focus-forming virus; mcs: multicloning site; IRES: internal ribosomal entry site; EGFP: enhanced green fluorescent protein. (B) Toxicity of HepG2 cells after 4-IPO exposure. Survival of EGFP+ HepG2 cells (%) after 24 and 48 h of incubation with 0, 2.9, 29, 290 or 1450 μM 4-IPO as measured by flow cytometry. The cells stably expressed the different CYP4B1 enzymes: h-S427: human wild-type (isoform 1); h-S427 insSer207: human wild-type with insSer207 (isoform 2), h-P427: human mutated p.S427P (artificial); h-P427 insSer207: human mutated p.S427P with insSer207 (artificial); r-P422: rabbit wild-type (isoform 1); r-P422 insSer202: rabbit with insSer202 (artificial). For each construct, the mean ± standard error of the mean (SEM) values are shown from at least three independent experiments. The transduction efficiency of HepG2 cells (not shown) was ≥90%, as determined by EGFP expression in flow cytometry. (C) Toxicity of T cells after 4-IPO exposure. Survival of EGFP+ primary human T cells (%) expressing the different CYP4B1 proteins [listed for (B)] after 24 and 48 h of incubation with increasing doses of 4-IPO as assessed by flow cytometry. For each construct, the mean ± SEM values are shown from at least three experiments.

doi:10.1371/journal.pone.0137110.g003

Otorhinolaryngology & head/neck surgery (ENT). This study including the documentation was approved by the ethics committee of the Heinrich-Heine-University of Düsseldorf (ethics No. 4687). After Ficoll-Hypaque density gradient centrifugation, outgrowth of >95% T cells was achieved by incubation of the mononuclear cells on immobilized CD3 and CD28 antibodies (BD Bioscience, San Jose, CA) in combination with IL-2 (100 IU/ml, Chiron, Marburg, Germany) in complete Iscove's Modified Dulbecco's Medium (Sigma-Aldrich, Deisenhofen, Germany) as described previously [28, 35, 36].

### Plasmid construction, lentivirus production, and cell transduction

The two lentiviral expression plasmids IRES-EGFP (used in Fig 3A and for Fig 4) and IRES-NeoR (Fig 5A) as well as expression constructs for the native rabbit CYP4B1 (NM\_001082103, isoform 1), the wild-type human CYP4B1 (NM\_000779, isoform 1), and the mutant human p.



**Fig 4. Bioactivation of 4-IPO to a stable NAC/NAL adduct.** The NADPH-dependent formation of NAC/NAL adducts by membranes isolated from HepG2 cells stably transduced with the different *CYP4B1* cDNAs r-P422: rabbit wild-type (isoform 1); h-S427: human wild-type (isoform 1); h-P427: human mutated p.S427P (artificial); h-S427 insSer207: human wild-type with insSer207 (isoform 2) after incubation in 50 nM 4-IPO for 20 minutes is shown. The adduct formation rate is shown as the mean  $\pm$  standard deviation (SD) of triplicate incubations.

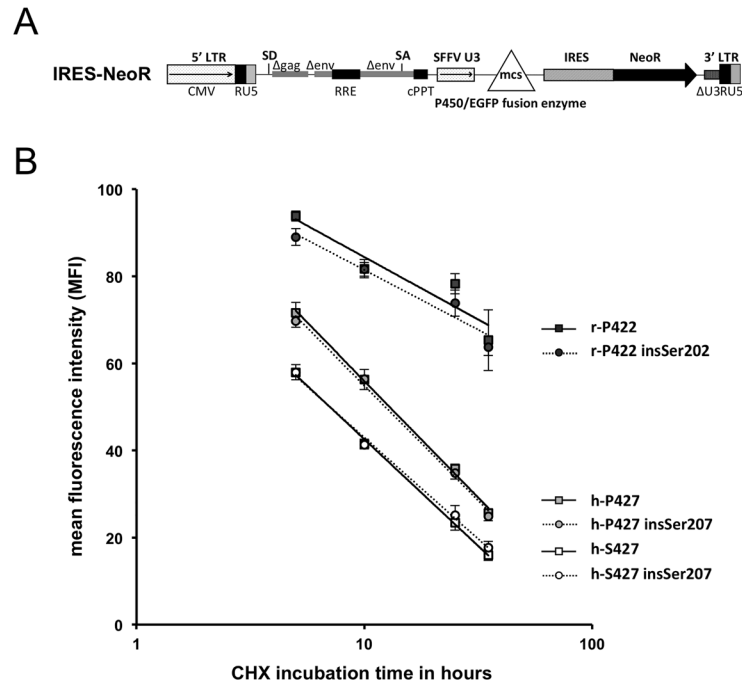
doi:10.1371/journal.pone.0137110.g004

S427P were described previously [28]. For insertion of serine at position 207 in the two human *CYP4B1* proteins (NM\_001099772: wild-type with insertion, isoform 2) and at position 202 in the rabbit enzyme, a commercial site-directed mutagenesis kit (Stratagene, La Jolla, CA) was used according to the manufacturer's recommendations with the following primers 5' - CGGCCTGGGCCACAGCAGGGACAGCAGCTACTACC (forward) and 5' -GGTAGTAGCTGCTGTCCCTGCTGTGGCCAGGCCG (reverse) for human *CYP4B1* and 5' -GGAGACAGTGCCTGAATCACAGCAGGGACAGCAGCTACTACGTG (forward) and 5' -CACGTAGTAGCTGCTGTCCCTGCTGTGATTCAGGCCACTGTCTCC (reverse) for rabbit *CYP4B1*. The cDNAs with a serine insertion, either with or without a 3'EGFP fusion, were then cloned into the IRES-NeoR and IRES-EGFP constructs, respectively, using the *NheI*-*Bam*HI restriction enzyme sites in the vectors. The final lentiviral expression constructs were verified by Sanger sequencing.

Vesicular stomatitis virus G (VSV-G) pseudotyped replication-deficient lentiviral particles were then produced after transfection into HEK293T cells and cryopreserved as described previously [28, 37]. Functional titers of the supernatants were assessed on HT1080 cells. HepG2 cells were transduced with either IRES-EGFP or IRES-NeoR vectors using 8  $\mu$ g/ml polybrene (Sigma-Aldrich) [28, 37]. Primary human T cells were transduced in the presence of 100 U/ml IL-2 on the recombinant fibronectin fragment CH-296 (Takara Shuzo, Otsu, Japan) as previously described [28, 35, 38, 39].

### Toxicity assay with 4-IPO

Five days after transduction of the HepG2 cells and four days after transduction of the T lymphocytes, cells were incubated in complete medium with increasing concentrations of 4-IPO (0, 2.9, 29, 290, and 1450  $\mu$ M). Cultures were harvested after 24 and 48 h, centrifuged, and then re-suspended in PBS with propidium iodide (PI; Sigma-Aldrich). Each sample was then analyzed by flow cytometry on a FACSCalibur (BD Biosciences, Heidelberg, Germany), and



**Fig 5. Protein stability of CYP4B1.** (A) Schematic outline of the lentiviral vector used for expression of CYP4B1. CMV: CMV promoter; SD: splice donor; LTR: long terminal repeat; SA: splice acceptor; RRE: Rev responsive element, cPPT: central polypurine binding tract; SFFV U3: U3 promoter of the spleen focus-forming virus; mcs: multicloning site; IRES: internal ribosomal entry site; neoR: neomycin phosphotransferase (nptII) resistance cDNA; EGFP: enhanced green fluorescent protein. (B) Protein half-life analysis of CYP4B1 isoforms in HepG2 cells. Protein half-life analysis was performed by adding 50 µg/ml cycloheximide (CHX) to the transduced and G418-selected HepG2 cells. The mean fluorescence intensity (MFI) of CYP4B1/EGFP fusion proteins was assessed at various time points by flow cytometry. For each construct, the mean ± SEM values are shown from at least three experiments.

doi:10.1371/journal.pone.0137110.g005

10,000 events were recorded. Live/dead cell analysis of EGFP+ cells was then performed using the CellQuest software as previously described [28]. Each experiment was performed independently at least three times.

### CYP4B1-dependent NAC/NAL adduct formation and quantitative analysis

*In vitro* bioactivation of 4-IPO by human CYP4B1 enzyme was assessed by incubating membrane preparations (1 mg) from transduced stably expressing HepG2 cells in the presence of substrate (1 mM), NADPH (1 mM), and both N-acetylcysteine (NAC) and N-acetyllysine (NAL) at 5 mM concentrations. Incubations were conducted at 37°C for 30 min, and adduct formation was quantified by a high performance liquid chromatography ultraviolet detector as previously described [40]

### Protein half-life analysis

For determination of the protein half-lives, we used 3'EGFP-fusion CYP4B1 proteins in the lentiviral IRES-NeoR vector (Fig 5A) as described previously [28]. Briefly, G418-resistant HepG2 cells were incubated with 50 µg/ml cycloheximide (CHX, Sigma-Aldrich), which inhibits eukaryotic protein translation. After 5, 10, 25, and 35 h, cells were harvested and fixed in 4% paraformaldehyde and, the mean fluorescence intensity (MFI) was assessed by flow cytometry.

The MFI values of cells treated with CHX were expressed as a percentage of the MFI of cells incubated in medium without CHX. Each experiment was performed independently at least three times.

## Modeling of human CYP4B1

Structural modeling of human CYP4B1 with insSer207 was performed using the Robetta server [41–43] that used the X-ray structure of a bacterial mimic of CYP2C9 (pdb id: 3QI8) as a template. Sequence identity between human CYP4B1 and the bacterial mimic of CYP2C9 was ~24%. A total of 5,000 models were generated followed by model clustering. The best CYP4B1 models were chosen based on the centers of the largest clusters and defined by having the lowest standard mean deviation value (between corresponding positions of C- $\alpha$  atoms of all residues) among all other models in a cluster.

## Results

### Evolution of the additional splice acceptor site in *CYP4B1*

To understand the molecular basis for alternative *CYP4B1* transcripts that do or do not encode insSer207, we first analyzed the splice donor and acceptor sites in intron 5 of human and rabbit *CYP4B1*. As shown in Fig 1A, the human genome harbors two splice acceptor sites in frame before exon 6, thus allowing the generation of two alternative transcripts, with or without serine at position 207, respectively. In contrast, intron 5 of the rabbit gene (Fig 1B) carries only a single splice acceptor site at the corresponding exon 6 boundary and therefore generates a single transcript without serine at amino acid position 202. Next, we aligned the genomic DNA of the corresponding *CYP4B1* intron 5 sequences of great apes, monkeys, and other mammals obtained from the ensembl genome database ([www.ensembl.org](http://www.ensembl.org)). The alignment in Fig 1C revealed that an insertion of a CAG triplet, generating the additional splice acceptor site, occurred during evolution at the level of *Hominoidae* (*Homo sapiens* and *apes*) and clearly distinguishes these species from old world (*rhesus macaque*) and new world (*marmoset*) monkeys. Other mammals such as rat, cow, and elephant also do not have the alternative splice acceptor site, thus making insSer207 specific for *Hominoidae*.

### Ratio of *CYP4B1* transcripts in healthy human tissues

To determine the usage of the two different splice acceptor sites of *CYP4B1* in human transcripts, a PCR with the primer pair shown in Fig 2A was performed on a commercial human cDNA library that was pooled from the mRNA of various organs obtained from 81 healthy donors. The PCR products were cloned into a high copy number plasmid, and the DNA inserts in bacterial colonies were isolated and sequenced. As shown in Fig 2B, 169 out of 220 (76.8%) bacterial clones contained insSer207, whereas 51 clones (23.2%) did not. Thus, the majority of translated human CYP4B1 contains insSer207.

### Functional analysis of human and rabbit CYP4B1 with/without insSer207/202

The enzymatic activation of 4-IPO is an established tool for assessing and comparing the activity of CYP4B1 orthologs. To evaluate the functional consequences of insSer207/202 on the activity of the CYP4B1 enzyme, PCR was used to introduce the CAG sequence, containing the additional splice acceptor site, into the human native and mutant p.S427P CYP4B1 enzymes at the intron 5–exon 6 boundary. We also introduced the serine insertion at the corresponding position 202 of the native rabbit CYP4B1 protein (insSer202). For stable expression in



mammalian cells, all six cDNAs were cloned into the lentiviral expression vector IRES-EGFP (Fig 3A), thus linking the expression of the *CYP4B1* cDNAs to EGFP as a marker gene for transduced cells [28]. Replication-deficient lentiviral particles pseudotyped with VSV-G were titered as described previously [28, 37] and used to transduce human HepG2 liver cells with the different constructs at similar viral particle per target cell, also called multiplicities-of-infection (MOIs) of approximately 5–7. The results demonstrated that HepG2 cells transduced with the control vector did not show any cell death in response to the increasing 4-IPO concentration after 24 and 48 h (Fig 3B), because bioactivation could not occur. Moreover, as shown previously [28], the native human *CYP4B1* enzyme was not capable of processing 4-IPO, while both the rabbit wild-type and the human p.S427P mutant *CYP4B1* induced clear cytotoxic effects on the cells. Finally, introduction of serine at corresponding positions into the wild-type rabbit or human p.S427P *CYP4B1* did not alter the high activities of both enzymes towards 4-IPO (Fig 3B).

To confirm these findings in non-transformed cells that do not express *CYP4B1* mRNA (Wiek & Hanenberg, unpublished), primary human T cells from healthy donors were isolated, transduced with the same vectors at similar MOIs, and then challenged with increasing doses of 4-IPO. After 24 and 48 h, the cells were harvested, and the survival of transduced EGFP+ T cells was analyzed by flow cytometry. As shown in Fig 3C, the insSer207/202 did not restore the negligible activity of the native human *CYP4B1* enzyme nor alter the strong metabolizing activities of the wild-type rabbit or the mutant human p.S427P *CYP4B1* proteins. Therefore identical to the results obtained in HepG2 cells, the serine insertion did not alter the enzymatic activity of neither the wild-type rabbit and human nor the mutant human *CYP4B1* protein.

#### 4-IPO induced adduct formation and cellular cytotoxicity

As survival of stably transduced human cells was used as the indicator for *CYP4B1* enzyme activity towards 4-IPO, we wanted to confirm that a direct correlation exists between the percentage of EGFP+ cells that survived after 4-IPO incubation and 4-IPO adduct formation in these cells. To this end, we assessed NAC/NAL adduct formation in membrane preparations from HepG2 cells stably transduced with the rabbit or one of the three human *CYP4B1* cDNAs. Just as 4-IPO was not toxic to cells expressing wild-type human *CYP4B1* (Fig 3B), membranes prepared from these cells did not catalyze the bioactivation of 4-IPO (Fig 4). However, when the serine in the meander region was reverted to the evolutionarily conserved proline at position 427 (p.Ser427Pro), 4-IPO was bioactivated and formed NAC/NL adducts (Fig 4). The very active rabbit enzyme served as a positive control. The results clearly demonstrated that both assays, cell survival and NAC/NAL adduct formation, are well suited for measuring the 4-IPO metabolizing activity of different *CYP4B1* enzymes. Importantly, both assays also showed that the insSer207 variant of human *CYP4B1* does not 'rescue' the loss of enzyme activity towards xenobiotics caused in large part by the presence of the meander region amino acid, Ser427.

#### Protein stability of human and rabbit *CYP4B1* with or without serine insertion

We have previously shown that the differences in the 4-IPO metabolizing activity of the wild-type rabbit and human mutant *CYP4B1* enzymes is caused, at least in part, by differences in protein half-life and expression level [28]. Given that Imaoka *et al.* suggested that the function of insSer207 may be to stabilize the native human *CYP4B1* enzyme [30], we next tested the effect of this insertion on the half-lives and expression levels of the native rabbit and human *CYP4B1* proteins. We also included the human *CYP4B1* p.S427 mutant with and without ins-Ser207 as additional controls.

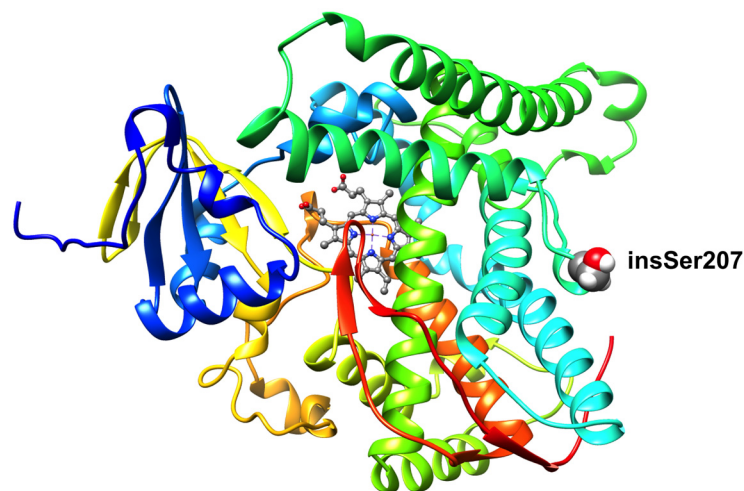
As no antibody is available that can recognize both the human and the rabbit proteins equally well, 3'EGFP-CYP4B1 fusion proteins were cloned into a lentiviral vector with an IRES-NeoR cassette (Fig 5A) as described previously [28]. HepG2 cells were transduced at similar MOIs of approximately 3 with the infectious lentiviral particles stably expressing the different fusion enzymes and then selected with G418 (Geneticin<sup>®</sup>). G418-resistant HepG2 cells were incubated with cycloheximide (CHX) and then harvested at several time points to semi-quantitatively assess the mean fluorescent intensities (MFIs) of the CYP4B1-EGFP fusion constructs by flow cytometry. Similarly to our previously reported findings [28], HepG2 cells expressing the native rabbit CYP4B1 showed the highest MFI (500 on the FL1 channel) of all EGFP-tagged proteins and the smallest decline during the observation period ( $t_{1/2} > 75$  h). These characteristics of the wild-type rabbit protein were not influenced by insSer202. The analysis further demonstrated that the corresponding insSer207 in the human protein also did not influence the half-life and expression level of either the native ( $t_{1/2} \sim 7$ h) or the mutant p.S427P ( $t_{1/2} \sim 13$ h) CYP4B1 protein (Fig 5B, pairs of solid versus hatched lines).

### Structural models of the serine insertions in human and rabbit CYP4B1

Finally, to better understand why insSer207 had no effect on the expression level, half-life, or enzymatic activity of the three CYP4B1 proteins, we generated a three-dimensional model for insSer207 in human CYP4B1. Because no crystal structure is available for any CYP4 protein, we used the Rosetta-based Robetta server to generate a homology model. The model shown in Fig 6 localizes insSer207 to a loop region, distant from the heme catalytic center, that is comprised by residues G198 to S210. Within this loop, the Ser207 side-chain is oriented away from the active site towards the soluble environment. These structural features are consistent with insSer207 having a minimal effect on human CYP4B1 structure and catalysis. Similar structural considerations also apply for the rabbit wild-type and human p.S427P CYP4B1 proteins (not shown).

### Discussion

Human CYP4B1 has been described as an enigmatic enzyme [5], in part due to the two human-specific serine residues that have been incorporated into the primary sequence during



**Fig 6. Structural model of human CYP4B1 with insSer207 located in a loop region from G198 to S210.** Cartoon representation of the structure of human CYP4B1. The side chain of insSer207 and the heme prosthetic group are shown as spacefilling and ball and stick representations, respectively. Carbon atoms are colored gray, oxygen atoms are colored red, nitrogen atoms are colored blue, and hydrogen atoms are colored white. The figure was generated using UCSF Chimera [43].

doi:10.1371/journal.pone.0137110.g006

evolution. One unanswered question regarding *CYP4B1* in the literature is the importance of insSer207 in human *CYP4B1* and its role in the function of the enzyme. Although not described in the initial cloning of the human *CYP4B1* cDNA [25], three groups subsequently described insSer207 as the major form of *CYP4B1* mRNA in human tissues [29–31]. Carr *et al.* reported that only this insertion was present in 5 out of 5 bacterial clones, but a more in-depth semi-quantification of the ratio of the two isoforms was not performed in any of the earlier studies. In the present work, after sequencing a library of exon 5-exon 6 amplified *CYP4B1* cDNA fragments, we demonstrated that the insSer207 is encoded by 76% of *CYP4B1* transcripts in humans. This insertion is the consequence of a CAG triplet that was introduced at the intron-exon junction of intron 5–6, thereby creating a second functional splice acceptor site for exon 6 in frame.

Imaoka *et al.* put forward the intriguing idea that insSer207 in native human *CYP4B1* might have occurred to restore the enzyme function by compensating for the negative consequences of the p.P427S alteration that had occurred in humans during evolution [30]. We therefore functionally tested the serine insertion in three different *CYP4B1* enzymes with well defined functional activities towards 4-IPO: i) the native rabbit protein, in which serine was artificially introduced at the corresponding amino acid position 202; ii) the native human *CYP4B1*; and iii) the mutant human *CYP4B1* p.S427P enzyme, in which the canonical proline present in other *CYP450*s at corresponding positions was re-introduced in place of serine. For the latter protein, we recently reported that this p.S427P amino acid change restored the metabolizing activities of the human protein for hallmark substrates of animal *CYP4B1* enzymes [28]. The results of the present study clearly demonstrate that insSer207 in the two human proteins and insSer202 in the rabbit enzyme have no functional impact on the bioactivation of 4-IPO in human primary cells and cell lines. In addition, the insertion did not alter the *CYP4B1* expression level or half-life when stably expressed in HepG2 cells. Importantly, in contrast to the hypothesis by Imaoka *et al.* [30], insSer207 did not restore the absent enzymatic activity of native human *CYP4B1* with serine at position 427 for 4-IPO. Thus, insSer207 in the human mutant p.S427P and insSer202 in the wild-type rabbit *CYP4B1* seem to be neutral modifications regarding the xenobiotic metabolizing capabilities of human and animal *CYP4B1*s.

The homology model for human *CYP4B1* mapped insSer207 to a peripheral loop region, and although this localization might support some local structural perturbation, insSer207 clearly does not affect any of the functional read-outs for the three pairs of enzymes studied here (Figs 3–5). In addition, the phylogenetic analysis demonstrated that during evolution, the insertion of a CAG repeat in intron 5 had already occurred at the stage of *hominoidae* and is therefore present in all great apes and humans (data on the *CYP4B1* locus in gibbons as lesser great apes are not available in public databases). As the p.P427S ‘variant’ is unique to humans and arose *later in evolution* than insSer207, the latter has no possible role in rescuing defective human *CYP4B1* enzyme activity.

For more than 25 years now, the role and function of the human *CYP4B1* has remained a mystery. Retrospectively, this uncertainty seems to be the consequence of a combination of factors, predominantly i) technical difficulties in reliably expressing the native human *CYP4B1* enzyme in human cells and thus in performing functional analyses, ii) contradictory results in heterologous expression systems such as insect cells and liver cells of transgenic mice, and iii) the absence of known functions for *CYP4B1* in mammals and lower organisms. However, in the last 2 years, we have made considerable progress towards achieving a better understanding of the structure–function relationships for human *CYP4B1*. Previously, we demonstrated that compared to the very active rabbit *CYP4B1* enzyme, the human protein lacks important amino acid determinants in the middle domain of the protein that can render even the human native protein with serine at position 427 able to actively process 4-IPO [28]. Here, we have

demonstrated that the two naturally occurring human *CYP4B1* isoforms, with/without ins-Ser207, have no 4-IPO bioactivating ability. Therefore, we can safely conclude that the native human *CYP4B1* enzyme lost its enzymatic activity against typical xenobiotic pneumotoxins during evolution and therefore is now an 'orphan' P450 enzyme in humans. However, it should be noted that it still remains unclear which functional role(s) the active *CYP4B1* homologs in great apes and other mammals might have, especially for endogenous substrates.

Finally, the conclusion, that the human *CYP4B1* protein, with or without the serine insertion, is unstable, non-functional, and thus redundant in humans fits well with the relatively high frequency of detrimental genetic polymorphisms found in two human studies [44, 45]. Lo-Guidance *et al.* detected five allelic variants of *CYP4B1* in a French Caucasian population with the second most frequent, *CYP4B1*\*2, present in 518 out of 2082 (24.9%) French-Canadian individuals. This allele represents a 2-bp deletion (c.881-2delAT) that results in a prematurely terminated protein [44]. Forty study participants (2%) carried this non-functional allele in the homozygous state [44]. Hiratsuka *et al.* reported that the frequency of a *CYP4B1* allele with the same 2-bp deletion is 33.3% in the Japanese population and that 18 out of 192 (9.4%) individuals are homozygous [45]. These high frequencies of other non-functional *CYP4B1* alleles support the theory that the gene is redundant in humans and therefore not subject to further genetic selection [46].

## Acknowledgments

We thank Edward J. Kelly, University of Washington, Seattle, WA, for critical discussions and Laney Weber, Houston, TX for editing of the manuscript.

## Author Contributions

Conceived and designed the experiments: EMS CW CMK AER HH. Performed the experiments: EMS CW OTP KR MF MG NL. Analyzed the data: EMS CW OTP KR CMK AER HH. Contributed reagents/materials/analysis tools: CAS CMK VYY. Wrote the paper: EMS CW OTP KR MF MG NL CAS CMK AER HH.

## References

1. Gotoh O. Evolution of cytochrome p450 genes from the viewpoint of genome informatics. *Biol Pharm Bull* 2012; 35(6):812–7. PMID: [22687468](#)
2. Omura T. Contribution of cytochrome P450 to the diversification of eukaryotic organisms. *Biotechnol Appl Biochem* 2013; 60(1):4–8. doi: [10.1002/bab.1099](#) PMID: [23586987](#)
3. Nelson DR. The cytochrome p450 homepage. *Hum Genomics* 2009; 4(1):59–65. PMID: [19951895](#)
4. Yu X, Cojocaru V, Wade RC. Conformational diversity and ligand tunnels of mammalian cytochrome P450s. *Biotechnol Appl Biochem* 2013; 60(1):134–45. doi: [10.1002/bab.1074](#) PMID: [23587001](#)
5. Baer BR, Rettie AE. *CYP4B1*: an enigmatic P450 at the interface between xenobiotic and endobiotic metabolism. *Drug Metab Rev* 2006; 38(3):451–76. PMID: [16877261](#)
6. Boyd MR. Role of metabolic activation in the pathogenesis of chemically induced pulmonary disease: mechanism of action of the lung-toxic furan, 4-ipomeanol. *Environ Health Perspect* 1976; 16:127–38. PMID: [1017416](#)
7. Serabjit-Singh CJ, Wolf CR, Philpot RM. The rabbit pulmonary monooxygenase system. *Immunochemical and biochemical characterization of enzyme components*. *J Biol Chem* 1979; 254(19):9901–7. PMID: [114516](#)
8. Wolf CR, Statham CN, McMenamin MG, Bend JR, Boyd MR, Philpot RM. The relationship between the catalytic activities of rabbit pulmonary cytochrome P-450 isozymes and the lung-specific toxicity of the furan derivative, 4-ipomeanol. *Mol Pharmacol* 1982; 22(3):738–44. PMID: [7155130](#)
9. Czerwinski M, McLemore TL, Philpot RM, Nhamburo PT, Korzekwa K, Gelboin HV, et al. Metabolic activation of 4-ipomeanol by complementary DNA-expressed human cytochromes P-450: evidence for species-specific metabolism. *Cancer Res* 1991; 51(17):4636–8. PMID: [1651809](#)

10. Verschoyle RD, Philpot RM, Wolf CR, Dinsdale D. CYP4B1 activates 4-ipomeanol in rat lung. *Toxicol Appl Pharmacol* 1993; 123(2):193–8. PMID: [8248926](#)
11. Wilson BJ, Yang DT, Boyd MR. Toxicity of mould-damaged sweet potatoes (*Ipomoea batatas*). *Nature* 1970; 227:521–2. PMID: [5428477](#)
12. Wilson BJ, Boyd MR, Harris TM, Yang DT. A lung oedema factor from mouldy sweet potatoes (*Ipomoea batatas*). *Nature* 1971; 231(5297):52–3. PMID: [4930474](#)
13. Boyd MR, Wilson BJ. Isolation and characterization of 4-ipomeanol, a lung-toxic furanoterpenoid produced by sweet potatoes (*Ipomoea batatas*). *J Agric Food Chem* 1972; 20(2):428–30. PMID: [5016626](#)
14. Boyd MR, Wilson BJ, Harris TM. Confirmation by chemical synthesis of the structure of 4-ipomeanol, a lung-toxic metabolite of the sweet potato, *Ipomoea batatas*. *Nature* 1972; 236(66):158–9.
15. Boyd MR, Burka LT, Wilson BJ. Distribution, excretion, and binding of radioactivity in the rat after intraperitoneal administration of the lung-toxic furan, [14C]4-ipomeanol. *Toxicol Appl Pharmacol* 1975; 32(1):147–57. PMID: [1135870](#)
16. Boyd MR, Burka LT. In vivo studies on the relationship between target organ alkylation and the pulmonary toxicity of a chemically reactive metabolite of 4-ipomeanol. *J Pharmacol Exp Ther* 1978; 207(3):687–97. PMID: [731424](#)
17. Doster AR, Mitchell FE, Farrell RL, Wilson BJ. Effects of 4-ipomeanol, a product from mold-damaged sweet potatoes, on the bovine lung. *Vet Pathol* 1978; 15(3):367–75. PMID: [685084](#)
18. Serajit-Singh CJ, Wolf CR, Philpot RM, Plopper CG. Cytochrome p-450: localization in rabbit lung. *Science* 1980; 207(4438):1469–70. PMID: [6767272](#)
19. Christian MC, Wittes RE, Leyland-Jones B, McLemore TL, Smith AC, Grieshaber CK, et al. 4-*Ipomeanol*: a novel investigational new drug for lung cancer. *J Natl Cancer Inst* 1989; 81(15):1133–45. PMID: [2664191](#)
20. Kasturi VK, Dearing MP, Piscitelli SC, Russell EK, Sladek GG, O'Neil K, et al. Phase I study of a five-day dose schedule of 4-*Ipomeanol* in patients with non-small cell lung cancer. *Clin Cancer Res* 1998; 4(9):2095–102. PMID: [9748125](#)
21. Lakhanpal S, Donehower RC, Rowinsky EK. Phase II study of 4-ipomeanol, a naturally occurring alkylating furan, in patients with advanced hepatocellular carcinoma. *Invest New Drugs* 2001; 19(1):69–76. PMID: [11291834](#)
22. Rowinsky EK, Noe DA, Ettinger DS, Christian MC, Lubejko BG, Fishman EK, et al. Phase I and pharmacological study of the pulmonary cytotoxin 4-ipomeanol on a single dose schedule in lung cancer patients: hepatotoxicity is dose limiting in humans. *Cancer Res* 1993; 53(8):1794–801. PMID: [8467498](#)
23. Robertson IG, Serajit-Singh C, Croft JE, Philpot RM. The relationship between increases in the hepatic content of cytochrome P-450, form 5, and in the metabolism of aromatic amines to mutagenic products following treatment of rabbits with phenobarbital. *Mol Pharmacol* 1983; 24(1):156–62. PMID: [6865923](#)
24. Smith PB, Tiano HF, Nesnow S, Boyd MR, Philpot RM, Langenbach R. 4-*Ipomeanol* and 2-aminoanthracene cytotoxicity in C3H/10T1/2 cells expressing rabbit cytochrome P450 4B1. *Biochem Pharmacol* 1995; 50(10):1567–75. PMID: [7503758](#)
25. Nhamburo PT, Gonzalez FJ, McBride OW, Gelboin HV, Kimura S. Identification of a new P450 expressed in human lung: complete cDNA sequence, cDNA-directed expression, and chromosome mapping. *Biochemistry* 1989; 28(20):8060–6. PMID: [2574990](#)
26. Zheng YM, Fisher MB, Yokotani N, Fujii-Kuriyama Y, Rettie AE. Identification of a meander region proline residue critical for heme binding to cytochrome P450: implications for the catalytic function of human CYP4B1. *Biochemistry* 1998; 37(37):12847–51. PMID: [9737862](#)
27. Yokotani N, Sogawa K, Matsubara S, Gotoh O, Kusunose E, Kusunose M, et al. cDNA cloning of cytochrome P-450 related to P-450p-2 from the cDNA library of human placenta. Gene structure and expression. *Eur J Biochem* 1990; 187(1):23–9. PMID: [2298205](#)
28. Wiek C, Schmidt EM, Rollecke K, Freund M, Nakano M, Kelly EJ, et al. Identification of amino acid determinants in CYP4B1 for optimal catalytic processing of 4-*Ipomeanol*. *Biochem J* 2015; 465(1):103–14. doi: [10.1042/BJ20140813](#) PMID: [25247810](#)
29. Bylund J, Finnstrom N, Oliw EH. Gene expression of a novel cytochrome P450 of the CYP4F subfamily in human seminal vesicles. *Biochem Biophys Res Commun* 1999; 261(1):169–74. PMID: [10405341](#)
30. Imaoka S, Hayashi K, Hiroi T, Yabusaki Y, Kamataki T, Funae Y. A transgenic mouse expressing human CYP4B1 in the liver. *Biochem Biophys Res Commun*. 2001; 284(3):757–62. PMID: [11396967](#)
31. Carr BA, Ramakanth S, Dannan GA, Yost GS. Characterization of pulmonary CYP4B2, specific catalyst of methyl oxidation of 3-methylindole. *Mol Pharmacol* 2003; 63(5):1137–47. PMID: [12695542](#)

32. Choudhary D, Jansson I, Stoilov I, Sarfarazi M, Schenkman JB. Expression patterns of mouse and human CYP orthologs (families 1–4) during development and in different adult tissues. *Arch Biochem Biophys* 2005; 436(1):50–61. PMID: [15752708](#)
33. Dutcher JS, Boyd MR. Species and strain differences in target organ alkylation and toxicity by 4-ipomeanol. Predictive value of covalent binding in studies of target organ toxicities by reactive metabolites. *Biochem Pharmacol* 1979; 28(23):3367–72. PMID: [119540](#)
34. Yost GS, Buckpitt AR, Roth RA, McLemore TL. Mechanisms of lung injury by systemically administered chemicals. *Toxicol Appl Pharmacol* 1989; 101(2):179–95. PMID: [2510367](#)
35. Pollok KE, Hanenberg H, Noblitt TW, Schroeder WL, Kato I, Emanuel D, et al. High-efficiency gene transfer into normal and adenosine deaminase-deficient T lymphocytes is mediated by transduction on recombinant fibronectin fragments. *J Virol* 1998; 72(6):4882–92. PMID: [9573255](#)
36. Hanenberg H, Batish SD, Pollok KE, Vieten L, Verlander PC, Leurs C, et al. Phenotypic correction of primary Fanconi anemia T cells from patients with retroviral vectors as a diagnostic tool. *Exp Hematol*. 2002; 30(5):410–20. PMID: [12031647](#)
37. Nakano M, Kelly EJ, Wiek C, Hanenberg H, Rettie AE. CYP4V2 in Bietti's crystalline dystrophy: ocular localization, metabolism of  $\omega$ -3 polyunsaturated fatty acids and functional deficit of the p.H331P variant. *Mol Pharmacol* 2012; 82(4):679–86. PMID: [22772592](#)
38. Hanenberg H, Xiao XL, Dilloo D, Hashino K, Kato I, Williams DA. Colocalization of retrovirus and target cells on specific fibronectin fragments increases genetic transduction of mammalian cells. *Nat Med* 1996; 2(8):876–82. PMID: [8705856](#)
39. Hanenberg H, Hashino K, Konishi H, Hock RA, Kato I, Williams DA. Optimization of fibronectin-assisted retroviral gene transfer into human CD34+ hematopoietic cells. *Hum Gene Ther* 1997; 8(18):2193–206. PMID: [9449373](#)
40. Baer BR, Rettie AE, Henne KR. Bioactivation of 4-ipomeanol by CYP4B1: adduct characterization and evidence for an enedial intermediate. *Chem Res Toxicol* 2005; 18(5):855–64. PMID: [15892579](#)
41. Chivian D, Baker D. Homology modeling using parametric alignment ensemble generation with consensus and energy-based model selection. *Nucleic Acids Res* 2006; 34(17):e112. PMID: [16971460](#)
42. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* 2004; 32(Web Server issue):W526–31. PMID: [15215442](#)
43. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 2004; 25(13):1605–12. PMID: [15264254](#)
44. Lo-Guidice JM, Allorge D, Cauffiez C, Chevalier D, Lafitte JJ, Lhermitte M, et al. Genetic polymorphism of the human cytochrome P450 CYP4B1: evidence for a non-functional allelic variant. *Pharmacogenetics* 2002; 12(5):367–74. PMID: [12142726](#)
45. Hiratsuka M, Nozawa H, Konno Y, Saito T, Konno S, Mizugaki M. Human CYP4B1 gene in the Japanese population analyzed by denaturing HPLC. *Drug Metab Pharmacokinet* 2004; 19(2):114–9. PMID: [15499177](#)
46. Guengerich FP, Munro AW. Unusual cytochrome p450 enzymes and reactions. *J Biol Chem* 2013; 288(24):17065–73. doi: [10.1074/jbc.R113.462275](#) PMID: [23632016](#)