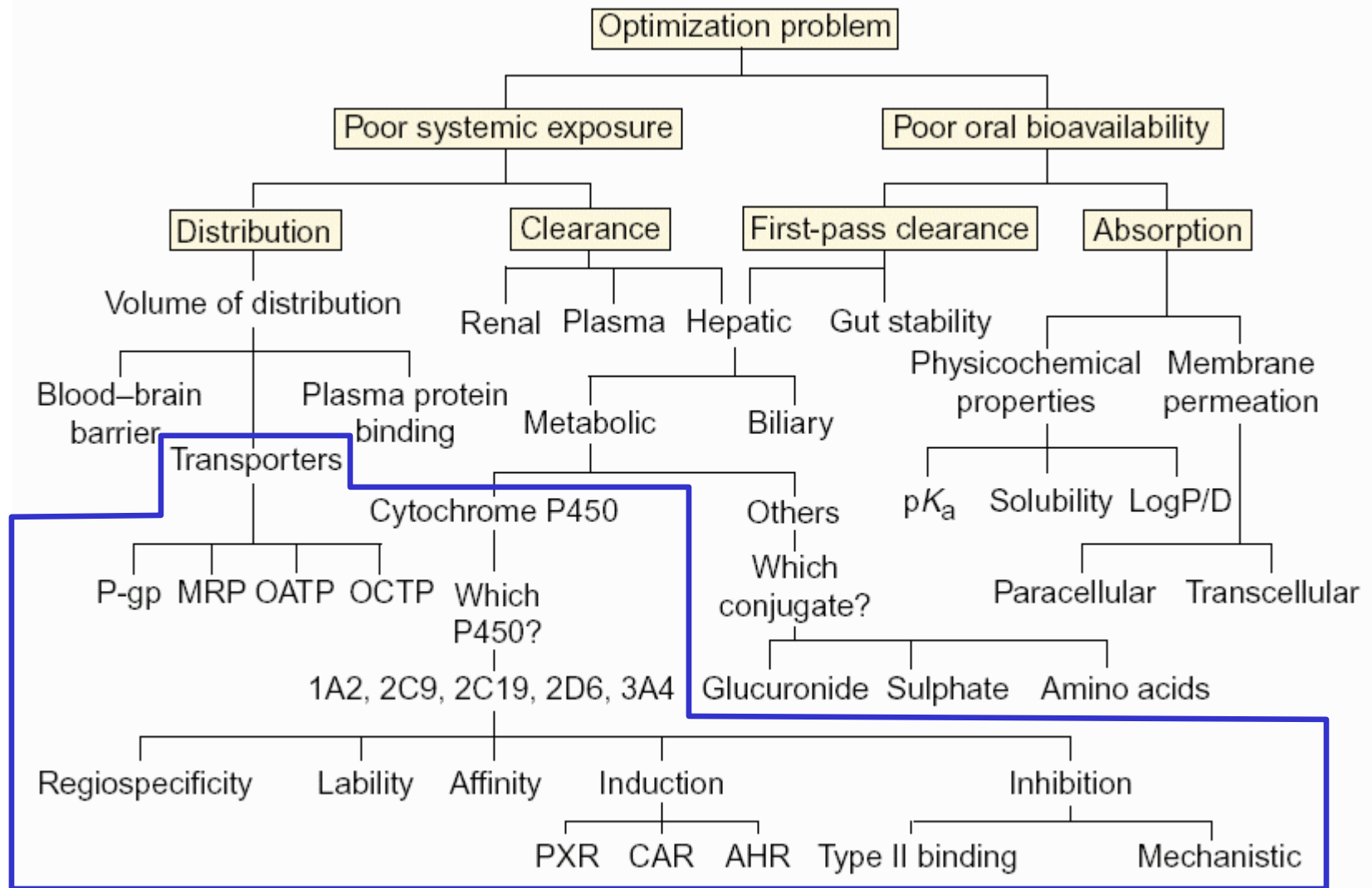
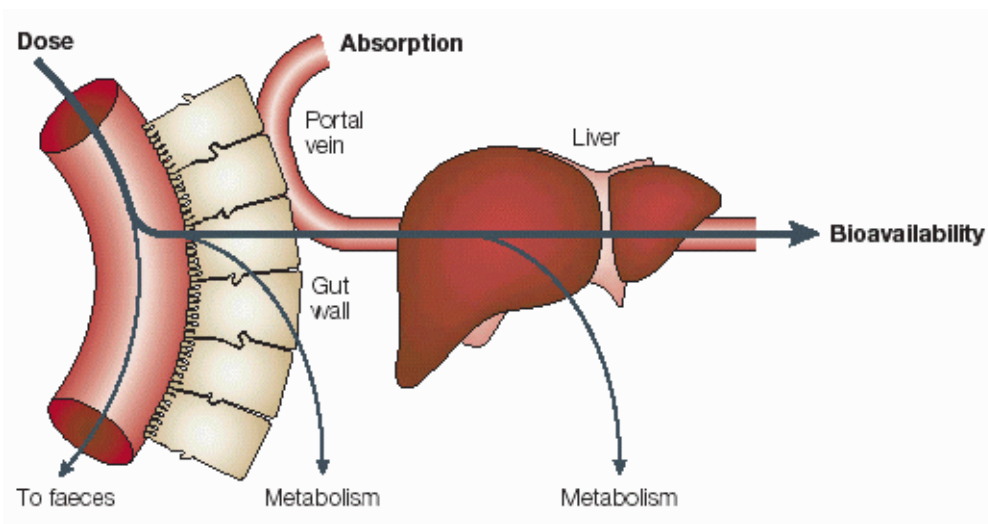


Cytochrome P450, Polymorphism, Transporters



Absorption and Metabolism

Nutrients as well as xenobiotics enter the blood circulation via the portal vein from the small intestine and reach the liver. Here, a variety of biochemical conversions of all substances is carried out.



Enzyme Systems That Metabolize Xenobiotics

Enzymatic System

Cytochrome P450

FAD-Monooxygenase

Monoamine Oxidase

Alcohol/Aldehyde Dehydrogenase

Epoxide Hydrolase

Gluthathione S-Transferase

Sulfotransferase

Acetyltransferase

Methyltransferase

Oxidoreductase

Xanthine Oxidase

Main Site of Location

Endoplasmatic Reticulum (5, 8)

Endoplasmatic Reticulum

Mitochondria (9)

Cytosol

Endoplasmatic Reticulum

Cytosol

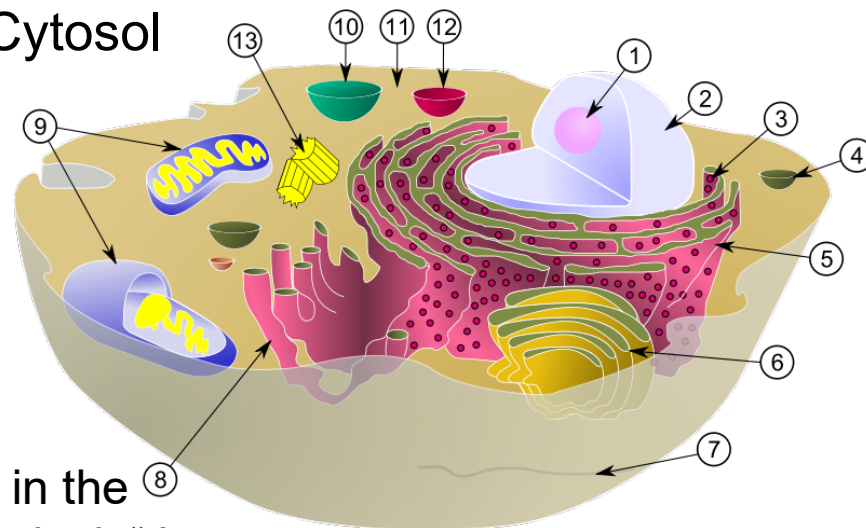
Cytosol

Cytosol

Cytosol

Cytosol

Cytosol



Lit: C. Ioannides „Cytochromes P450 in the Metabolism and Bioactivation of Chemicals“ in

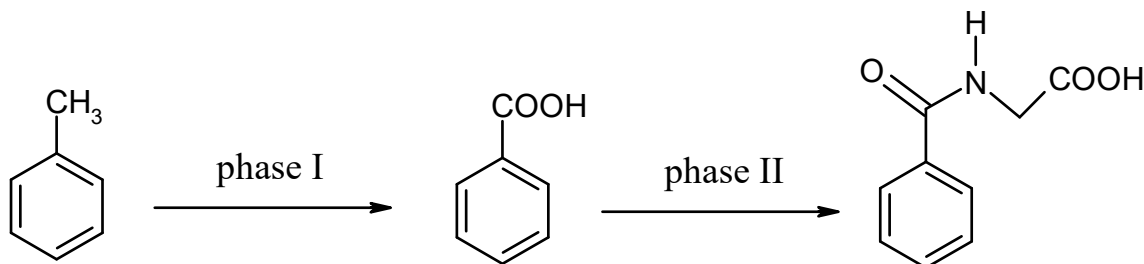
Chemistry and Molecular Aspects of Drug Design and Action,
Eds. E.A. Rekká, P.N. Kourounakis, CRC Press, Boca Raton, FL, 2008.

Picture: Wikipedia

Cytochrome P450 Metabolism (I)

First reactions: *First pass effect*

predominately lipophilic or heavy (MW >500) compounds are metabolized excessively, whereby they become more hydrophilic and thus easier to excret.



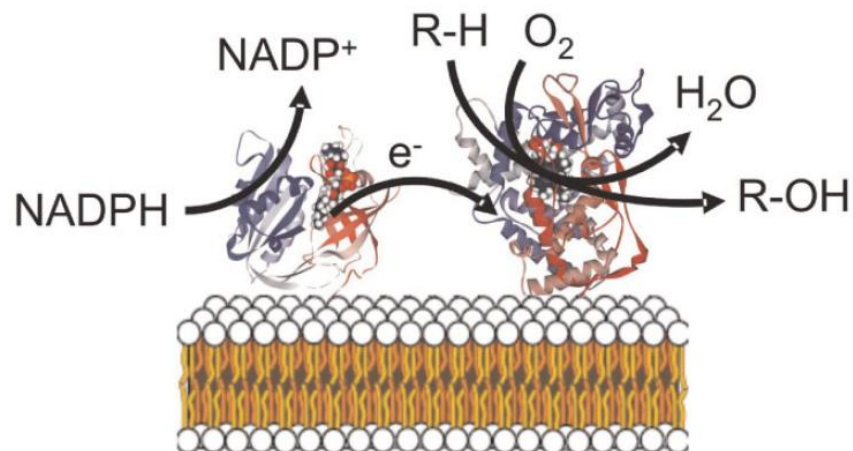
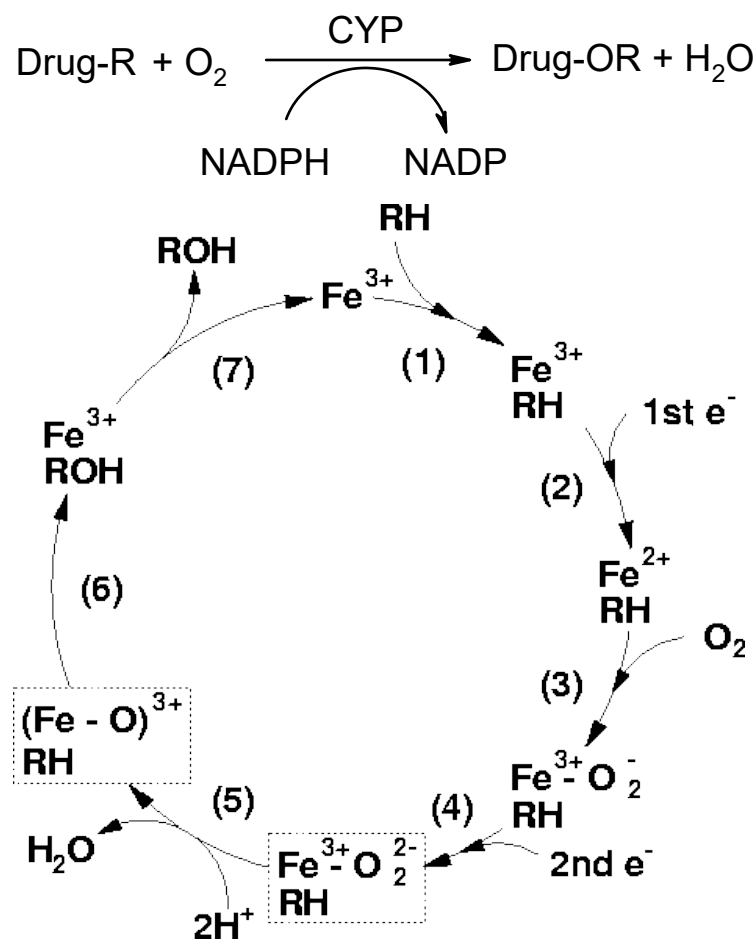
For the reactions comprising Phase I mainly the group of Cytochrome P450 enzymes (CYP) is responsible.

Usually substances are oxidized (formal addition of oxygen; redox reaction), however reduction and further chemical reactions (depending on the substrate) have been observed.

→ difficult to predict!

Cytochrome P450 Metabolism (II)

This mono-oxygenation of the substrates occurs in a catalytic cycle mediated by a hemoglobin-iron (Fe)



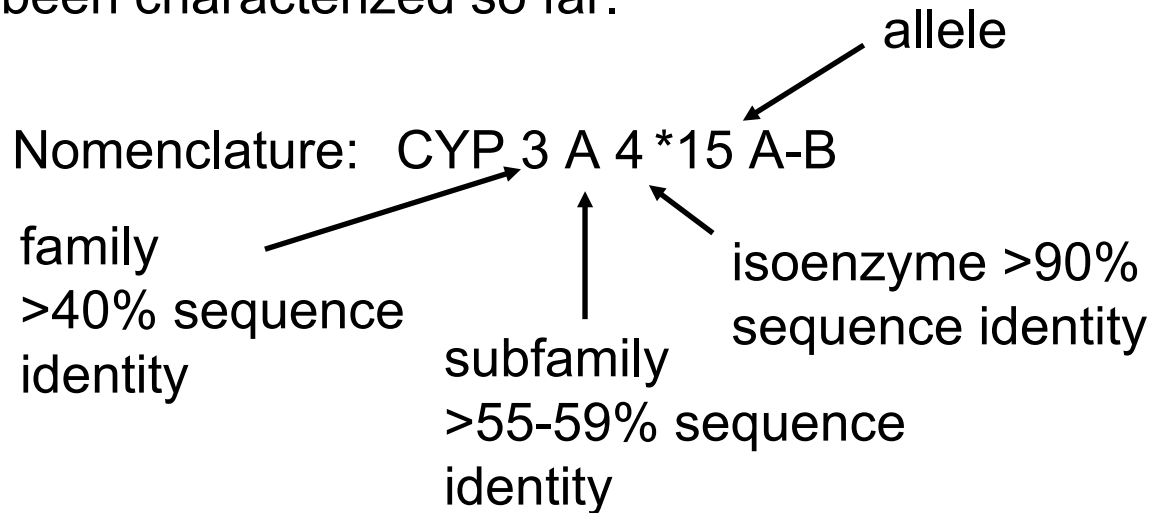
The electrons are provided by the cytochrome reductase

Cytochrome P450 Metabolism (III)

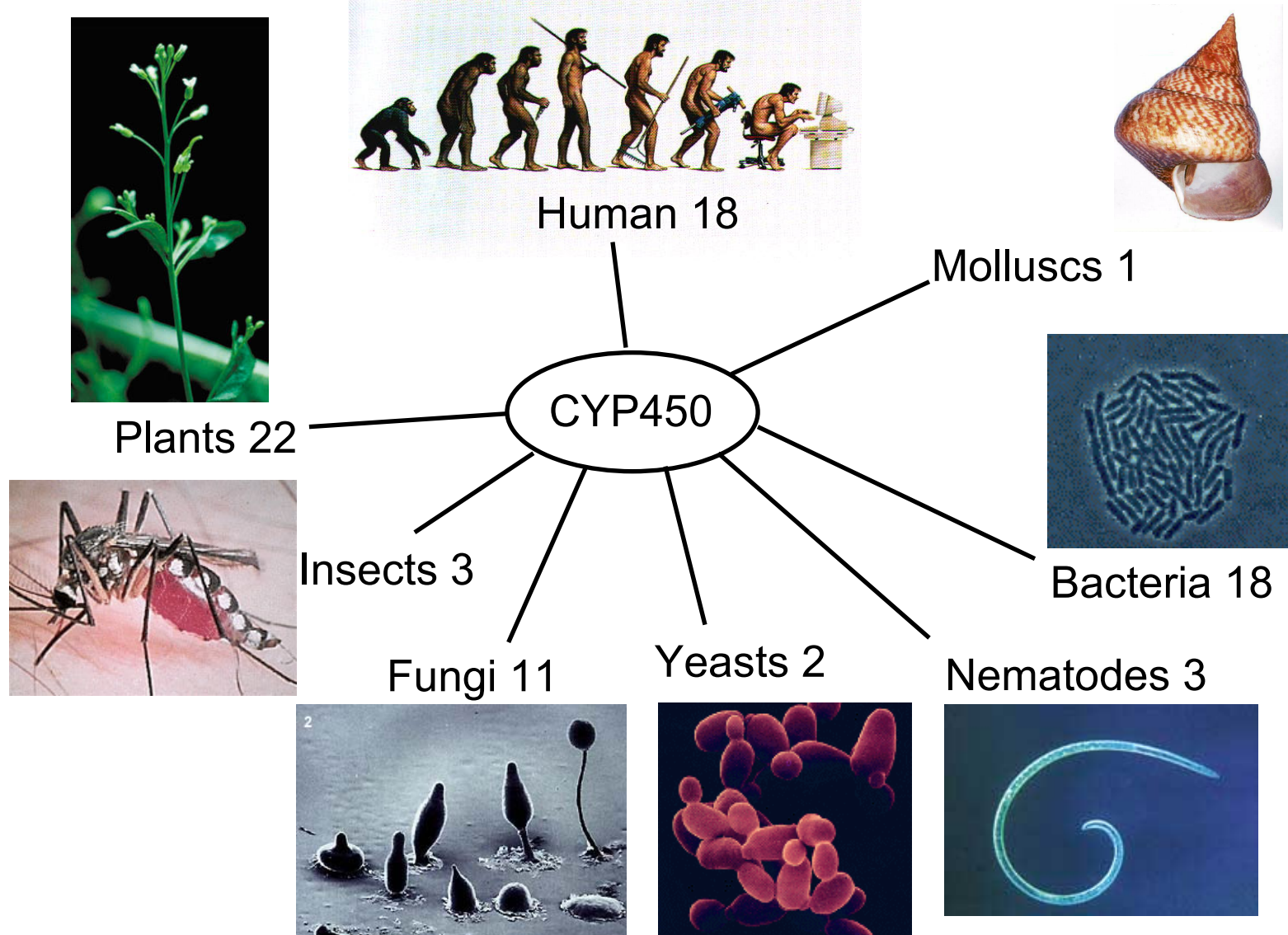
The cytochrome enzymes that account for the metabolism are predominately mono-oxygenases that evolved from enzymes for steroid and fatty acid synthesis. → cellular membranes

Sequence conservation is pretty low among CYPs in comparison to other enzymes.

In human 18 CYP-families containing about 58 isoforms have been characterized so far.



Cytochrome P450 Gene families



Human cytochrome P450 family

From the super-family of the cytochromes, the following families have been found in human:

CYP 1-5, 7, 8, 11, 17, 19-21, 24, 26, 27, 39, 46, 51

CYP 1, 2A, 2B, 2C, 2D, 2E, 2F, 3 metabolisms of xenobiotics

CYP 2G1, 7, 8B1, 11, 17, 19, 21, 27A1, 46, 51 steroid
metabolisms

CYP 2J2, 4, 5, 8A1 fatty acids metabolisms

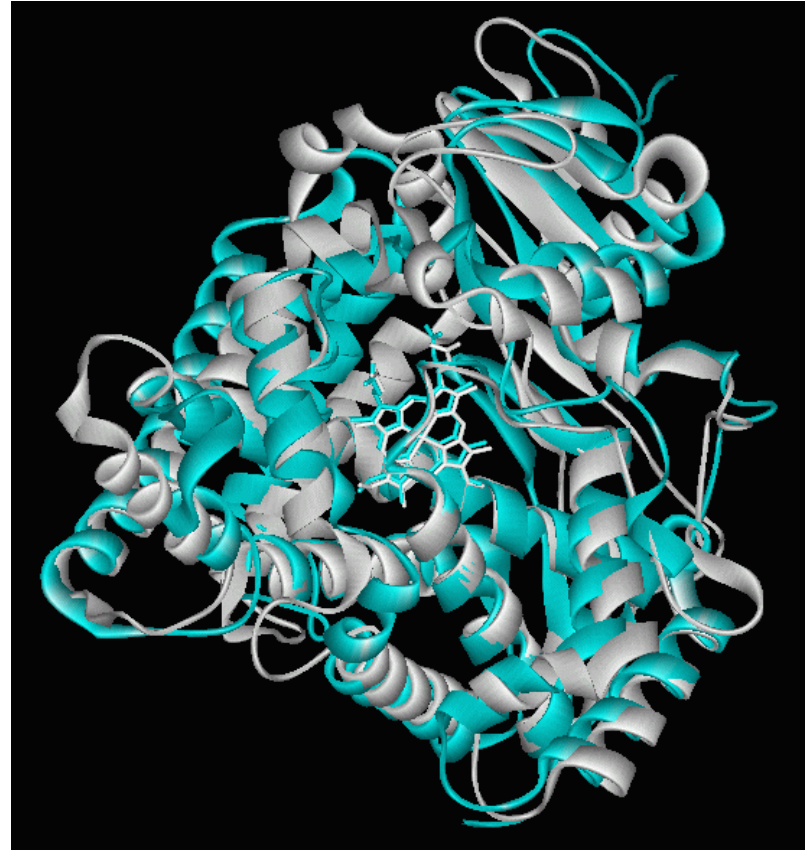
CYP 24 (vitamine D), 26 (retinoic acid), 27B1 (vitamine D), ...
synthesis

Cytochrome P450 Enzymes (I)

Despite the low sequence identity of CYPs from different species, the overall tertiary structure is conserved, esp. in the active center. In the outer regions, however, strong deviations occur.

Nevertheless, substrate and product specificity is governed by mutations in the binding pocket.

Superposition of human *h*CYP 2C9 (1OG5.pdb) and CYP 450 BM3 (2BMH.pdb) *Bacillus megaterium*



In contrast to bacterial CYPs, *mammalian* CYPs are typically partially embedded in the membrane. Lipophilic substrates can diffuse within the membrane to the CYP.

Cytochrome P450 Enzymes (II)

flavin monooxygenase isoenzyme (FMO)

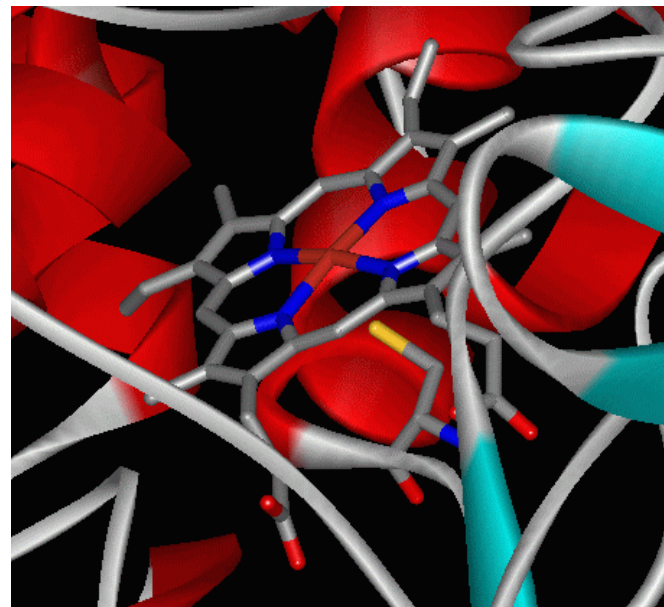
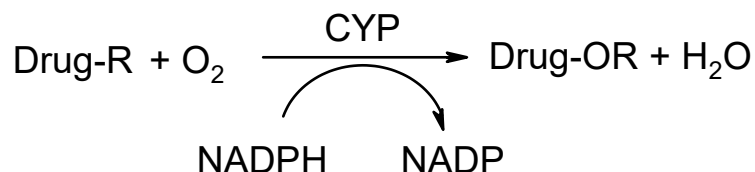
monoamine dehydrogenase (MAO)

aldo-keto reductase (AKR)

alcohol dehydrogenase

aldehyde oxidase

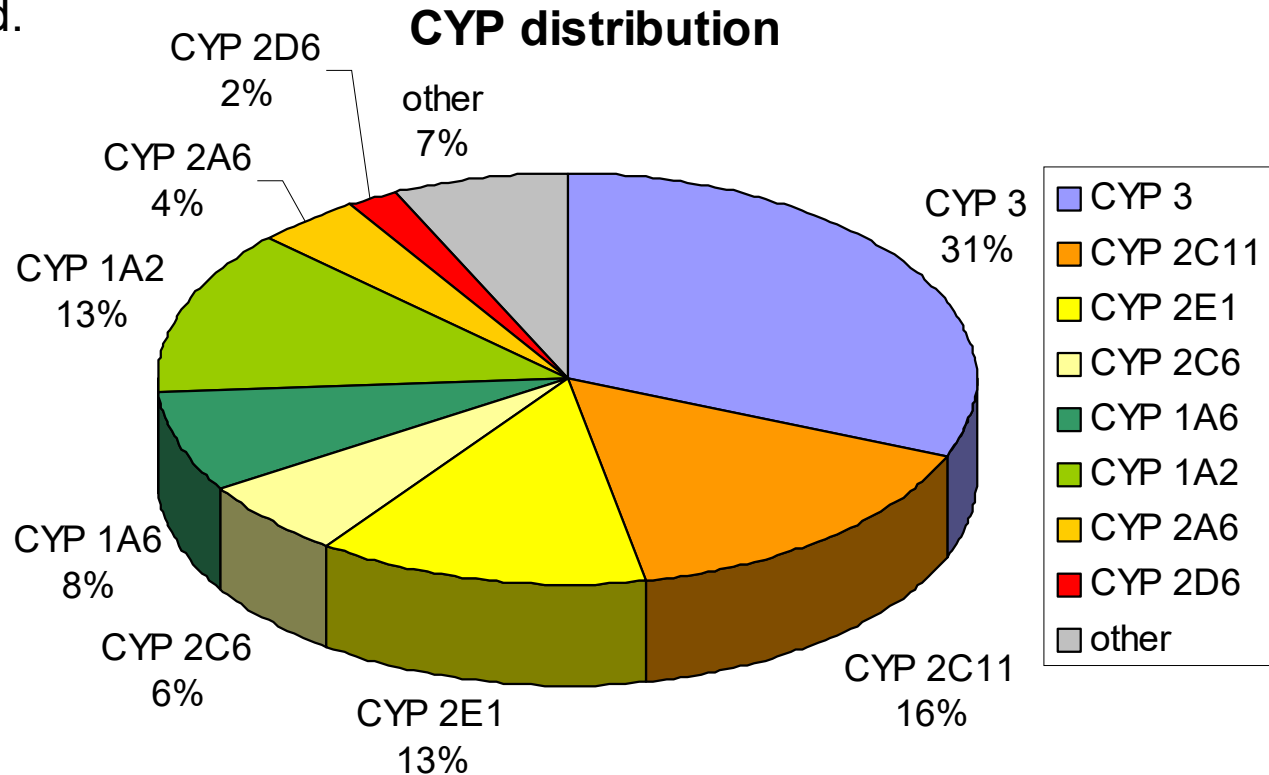
} Further
phase I
enzymes



Cytochrome P450 Enzymes (III)

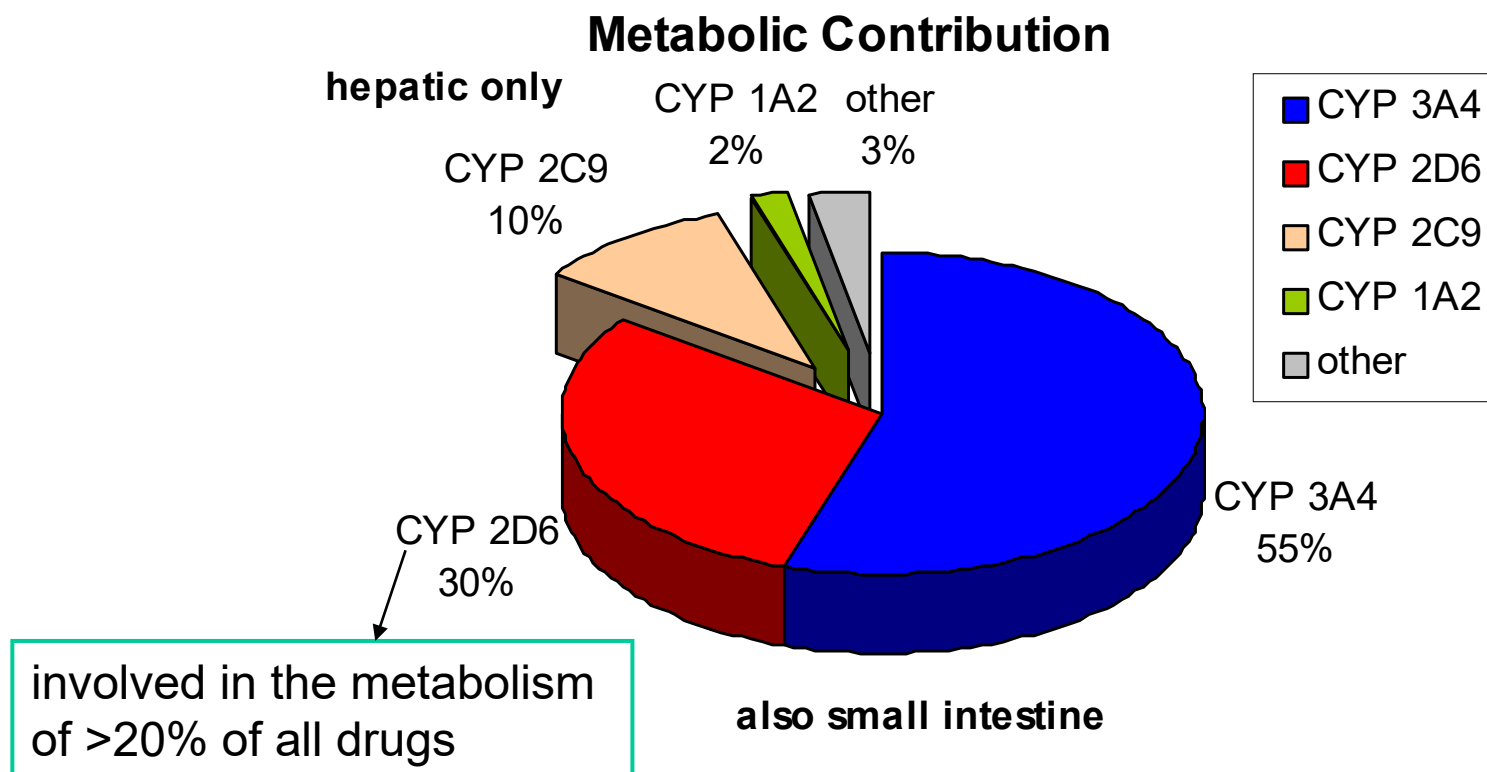
The prevailing amount of CYPs is present in the liver, however, certain CYPs are also expressed in cells of the intestine wall, lung, brain, heart, eye, mammal gland, and adrenal cortex.

The *mammalian* CYPs are closely attached to the membrane of the endoplasmatic reticulum via a membrane anchor, or are partially embedded.



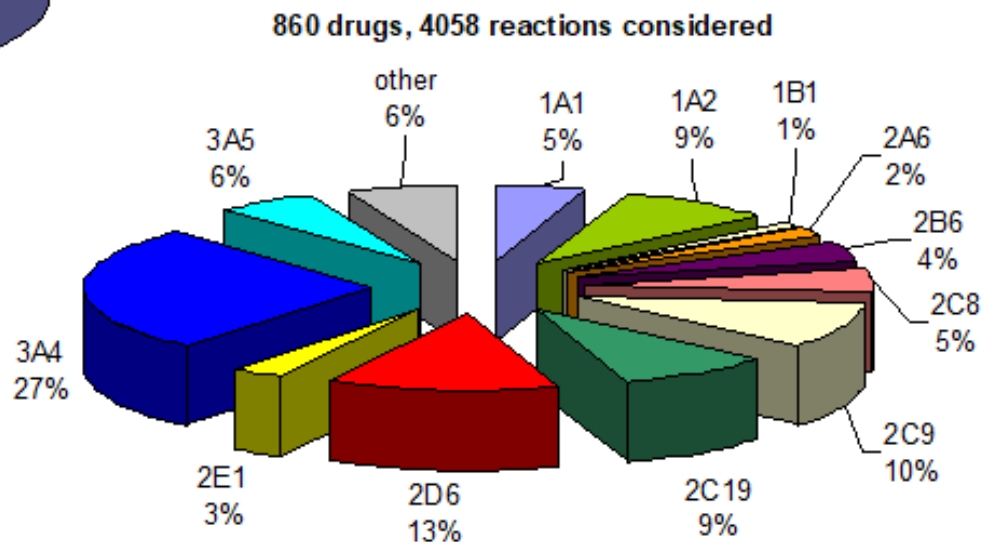
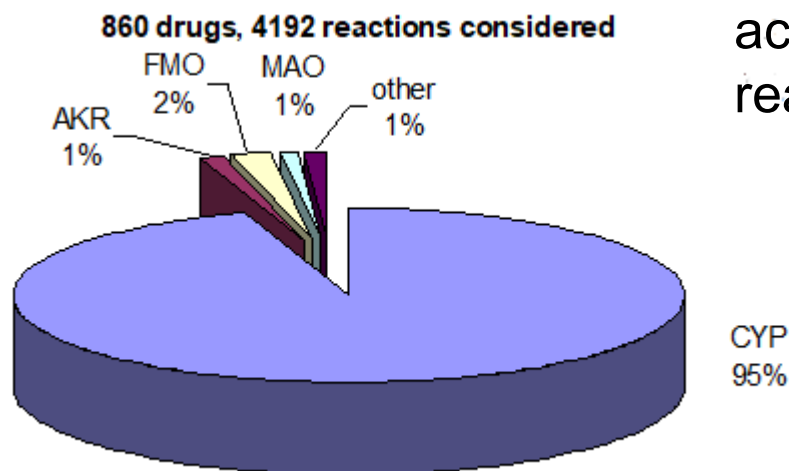
Cytochrome P450 Enzymes (IV)

The metabolism of endogenous substances (xenobiotics) is carried out predominately by CYP 3A4, CYP 2D6, and CYP 2C9.



Cytochrome P450 Enzymes (V)

Human oxidoreductases participating in the metabolism of drugs according to the number of chemical reactions



Lit. Rendic & Guengerich *Chem.Res.Toxicol.* **28** (2015) 39

Substrate specificity of CYPs (I)

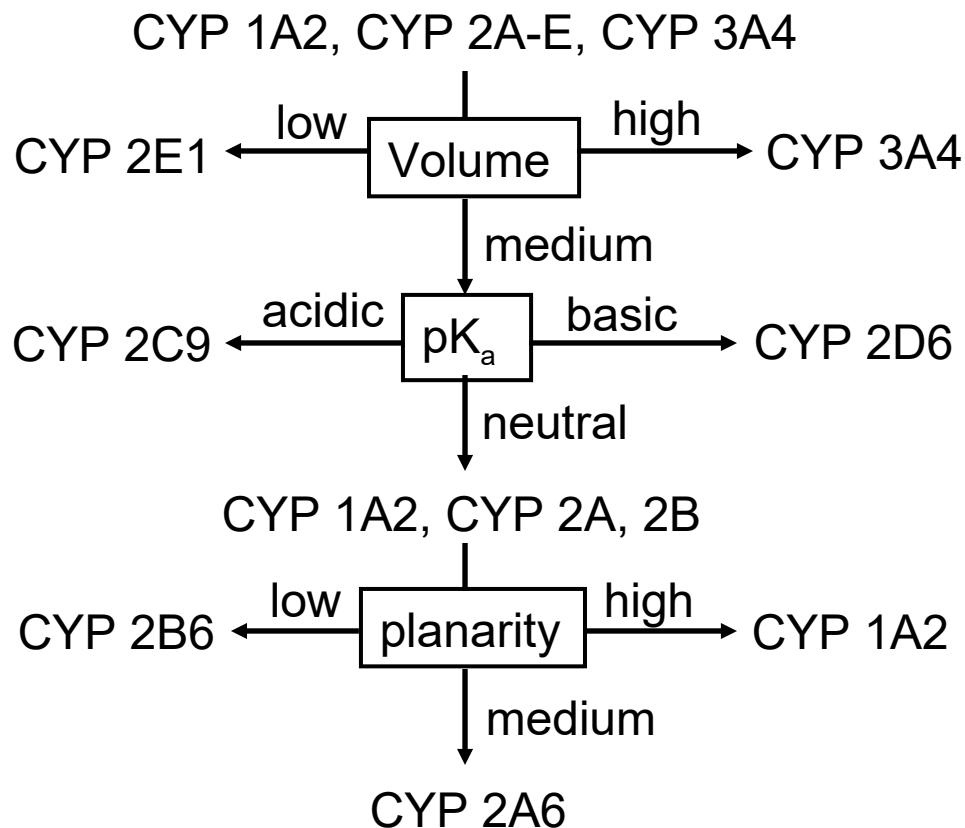
Specific substrates of certain human CYPs

CYP 1A2	verapamil, imipramine, amitryptiline, caffeine (arylamine <i>N</i> -oxidation)
CYP 2A6	nicotine
CYP 2B6	cyclophosphamid
CYP 2C9	diclofenac, naproxen, piroxicam, warfarin
CYP 2C19	diazepam, omeprazole, propranolol
CYP 2D6	amitryptiline, captopril, codeine, mianserin, chlorpromazine
CYP 2E1	dapsone, ethanol, halothane, paracetamol
CYP 3A4	alprazolam, cisapride, terfenadine, ...

Often substances are metabolized by multiple CYPs

Substrate specificity of CYPs (II)

„hand made“ decision tree for human P450 substrates

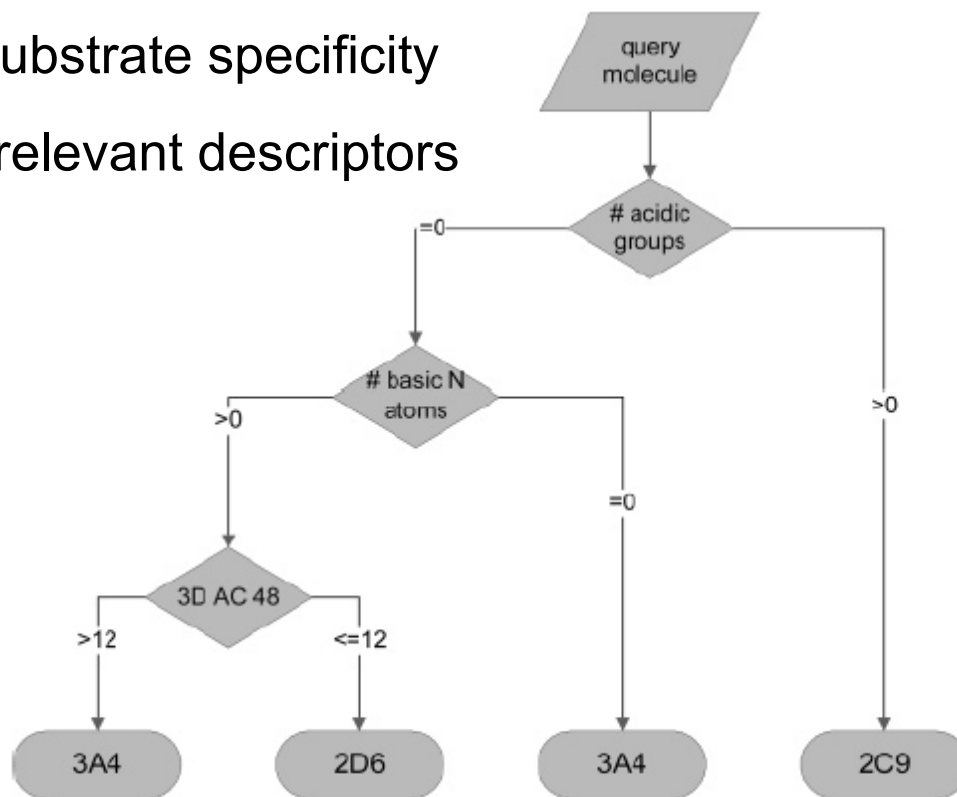


Lit: D.F.V. Lewis *Biochem. Pharmacol.* **60** (2000) 293

Prediction Models for Cytochrome P450 Metabolism (I)

Decision Tree for substrate specificity

→ Identification of relevant descriptors



Lit. L.Terfloth et al. *J.Chem.Inf.Model.* **47** (2007) 1688-1701.

Major public source of experimental data:
S.Rendic *Drug Metabol.Rev.* **34** (2002) 83-448.

Prediction Models for Cytochrome P450 Metabolism (II)

Qualitative prediction of metabolism for specific CYPs:

Binary classification into substrates / non-substrates

inhibitors / non-inhibitors

Problems: partial overlap of inhibitors and non-substrates

variability of data sets (how much of a non-substrate is metabolized?), unbalanced data sets (one class dominating)

Used machine learning algorithms: decision trees, neural networks, support vector machines, *k*-nearest neighbor, naïve Bayes

Lit. C.W.Yap & Y.Z.Chen *J.Chem.Inf.Model.* **45** (2005) 982-992.

J.M.Kriegel et al. *QSAR Comb.Sci.* **24** (2005) 491-502.

P.S.Bazeley et al. *J.Chem.Inf.Model.* **46** (2006) 2698-2708.

B.F.Jensen et al. *J.Med.Chem.* **50** (2007) 501-511.

M.Carbon-Mangels & M.C.Hutter. *J.Mol.Inf.* **30** (2011) 885-895.

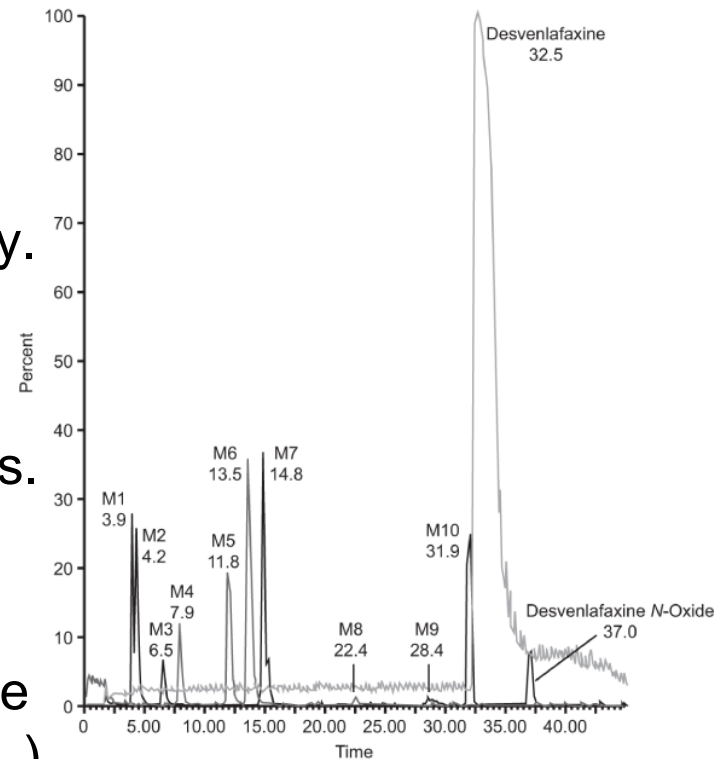
Cytochrome P450 Metabolism (IV)

During pre-clinical development it is of importance to characterize also the metabolic products of drugs since these might be toxic themselves.

Experimentally, the according (human) CYP-enzymes are expressed in *E. coli*, and the conversion is monitored by gas chromatography and mass spectroscopy.

This allows the selective determination of metabolites by single cytochrome P450 enzymes and their genetic variants.

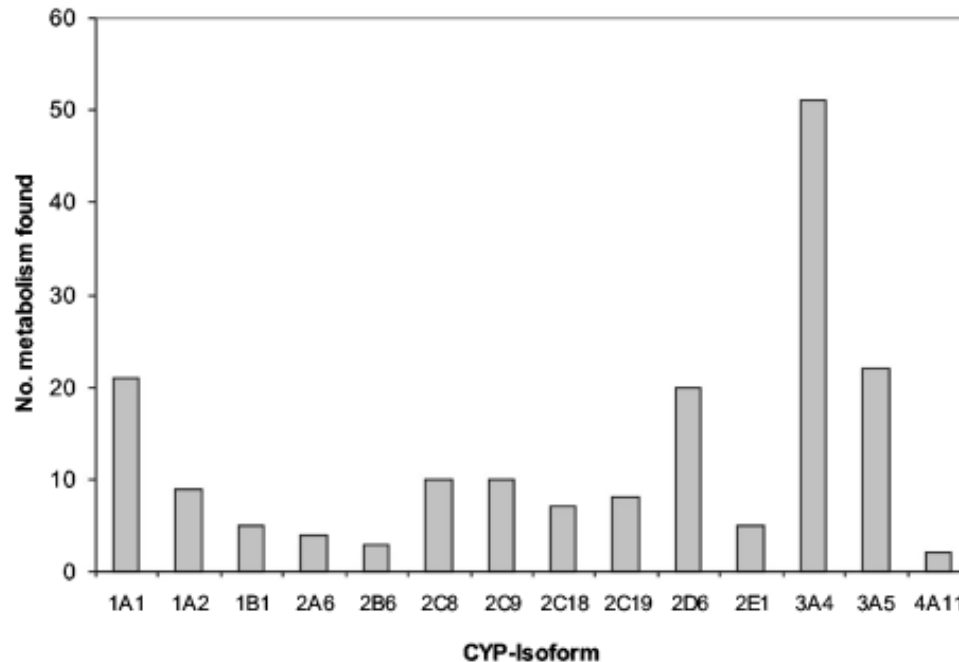
The results are compare to corresponding *in vivo* results from animals in order to chose the appropriate animal model (mouse, dog, guinea pig,...).



Lit. K.Schroer, M.Kittelmann, S.Lütz *Biotechnol. & Bioengin.* **106** (2010) 699.

Cytochrome P450 Metabolism (V)

The most prominent CYP-Enzymes during pre-clinical development used for generating metabolites



Number of metabolism events found (60 compounds tested)

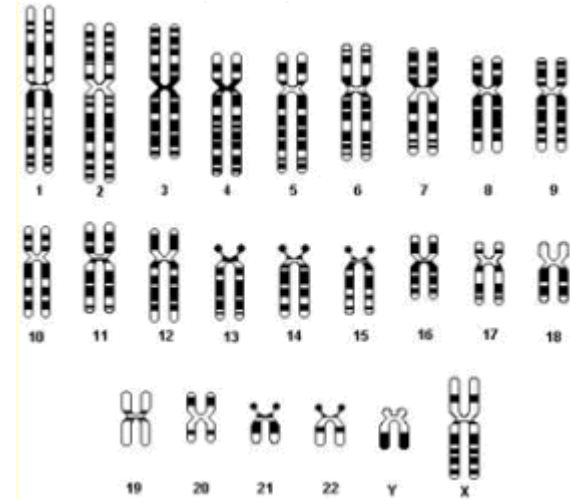
Lit. K.Schroer, M.Kittelmann, S.Lütz *Biotechnol. & Bioengin.* **106** (2010) 699.

Cytochrome P450 polymorphism

„Every human differs (more or less)“

The phenotype can be distinguished by the actual activity or the amount of the expressed CYP enzyme.

The genotype, however, is determined by the individual DNA sequence. Human: two sets of chromosomes (diploid)



That means: The same genotype enables different phenotypes

Depending on the metabolic activity, three major categories of metabolizers are separated: *extensive metabolizer* (normal), *poor metabolizer*, and *ultra-rapid metabolizer* (increased metabolism of xenobiotics)

Lit: K. Nagata et al. *Drug Metabol. Pharmacokin* **3** (2002) 167

Single Nucleotide Polymorphism (SNP)

SNPs are differences of single bases in the DNA that can be observed between individuals in a population.

Alleles occurring in at least 1% of the population are defined as polymorphism, which means that these genotypes occur regularly without causing diseases.

Conversely, differences in the genome that occur in less than 1% are referred to as mutations.

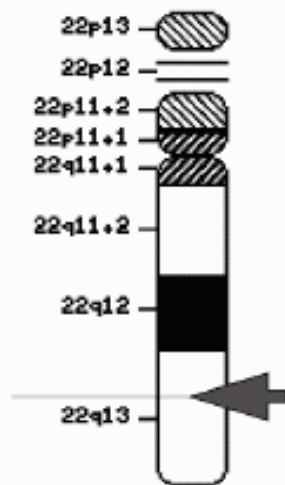
In the case of rare inherited diseases, typically mutations in the coding region of DNA sequences are observed.

Lit: A.D. Rose *Nature* **405** (2000) 857.

CYP 2D6 Polymorphism (I)

The polymorphism of CYP 2D6 (debrisoquine 4-hydroxylase) has been studied in great detail, as metabolic differences have first been described for debrisoquine and sparteine (antipsychotics).

See: D.B.Goldstein et al. *Nature Rev. Genetics* 4 (2003) 937.



CYP2D6*1 X N ↑ : Gene duplication
CYP2D6*2 X N ↑ : Gene duplication (N=2, 3, 4, 5 or 13)
CYP2D6*4X 2 : Gene duplication
CYP2D6*5 : Gene deletion
CYP2D6*13 : CYP2D7P/CYP2D6 hybrid, Exon1 CYP2D7P.E2-9 CYP2D6, frame shift
CYP2D6*16 : CYP2D7P/CYP2D6 hybrid, Exon1-7 CYP2D7P-related.E8-9 CYP2D6, Frame shift
CYP2D6*35 X 2 ↑ : Gene duplication
CYP2D6*36 ↓ : Conversion to CYP2D7 in exon 9

localized on chromosome 22

Of the 75 alleles, 26 exprime CYP2D6 proteines

see <http://www.imm.ki.se/CYPalleles/cyp2d6.htm>

CYP 2D6 Polymorphisms (II)

Designation	Characteristic mutation(s)	Enzyme activity	Allelic frequency (%)
<i>CYP2D6*1</i>	Wild type	Normal	
<i>CYP2D6*2</i>	G ₁₇₄₉ C, C ₂₉₃₈ T, G ₄₂₆₈ C substitutions	Normal	30
<i>CYP2D6*3</i>	A ₂₆₃₇ deletion	Deficient	2
<i>CYP2D6*4</i>	G ₁₉₃₄ A substitution	Deficient	22
<i>CYP2D6*5</i>	Gene deletion	Deficient	2
<i>CYP2D6*6</i>	T ₁₇₉₅ deletion	Deficient	2
<i>CYP2D6*7</i>	A ₃₀₂₃ C substitution	Deficient	0.1
<i>CYP2D6*8</i>	G ₁₈₄₆ T substitution	Deficient	0.1
<i>CYP2D6*9</i>	(A ₂₇₀₁ -A ₂₇₀₃) or (G ₂₇₀₂ -A ₂₇₀₄) deletion	Decreased	1.5
<i>CYP2D6*10</i>	C ₁₈₈ T, G ₁₇₄₉ C, G ₄₂₆₈ C substitutions	Decreased	1.5
<i>CYP2D6*11</i>	G ₉₇₁ C substitution	Deficient	0.1
<i>CYP2D6*12</i>	G ₂₁₂ A substitution	Deficient	0.1
<i>CYP2D6*13</i>	Hybrid: 2D7 exon 1, 2D6 exons 2-9	Deficient	0.1
<i>CYP2D6*14</i>	G ₁₈₄₆ A substitution	Deficient	0.1
<i>CYP2D6*15</i>	T ₂₂₆ insertion	Deficient	0.1
<i>CYP2D6*16</i>	Hybrid: 2D7 exons 1-7, 2D6 exons 8-9	Deficient	0.1
<i>CYP2D6*1</i> × 2	Gene duplication	Increased	1
<i>CYP2D6*2</i> × 2	Gene duplication	Increased	1.5
<i>CYP2D6*4</i> × 2	Gene duplication	Deficient	0.5

Lit: J. van der Weide et al. *Ann. Clin. Biochem* **36** (1999) 722

CYP 2D6 Polymorphism (III)

MGLEALVPLAVIVAIIFLLLVDLMHRRQRWAARYPPGPLPLPGLGNLLHVDFQNTPTCYCFDQ

poor debrisoquine metabolism S

R impaired mechanism of sparteine

LRRRFQDVFSLQLAWTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILGFGPRSQGVF

poor debrisoquine metabolism I

LARYGPAWREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDK

poor debrisoquine metabolism R

AVSNVIASLTCGRRFEYDDPRFLRLDLAQEGLKEESGFLREVLNAVVPVLLHIPALAGKV

LRFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEAKAKGNPESFNDENLRIVVA

missing in CYP2D6*9 allele

DLFSAGMVTSTTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI

P loss of activity in CYP2D6*7

HEVQRFGDIVPLGMTHMTSRDIEVQGFRI PKGTTLITNLSSVLKDEAVWEKPFRFHPEHF

LDAQGHFVKPEAFLPFSAGRRACLGEPLARMELFLFFTSLQLQHFSSVPTGQPRPSHHGV

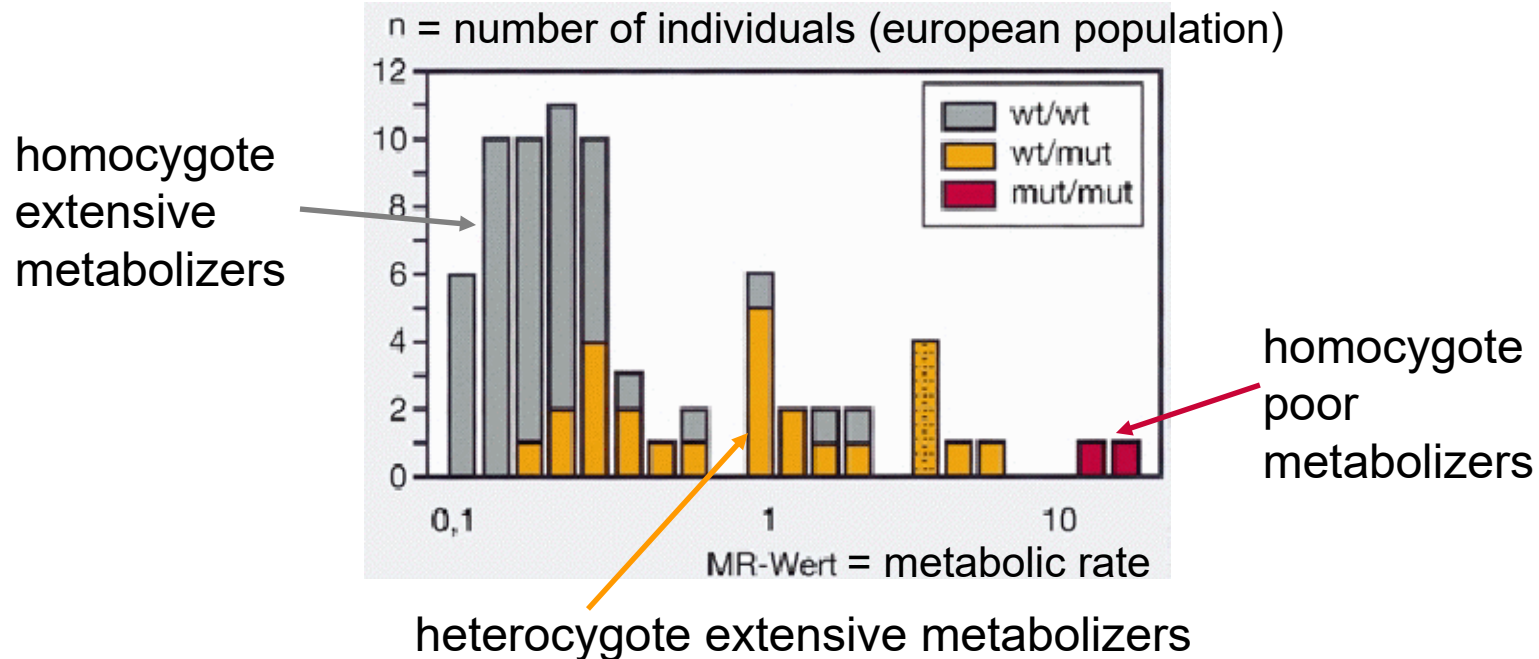
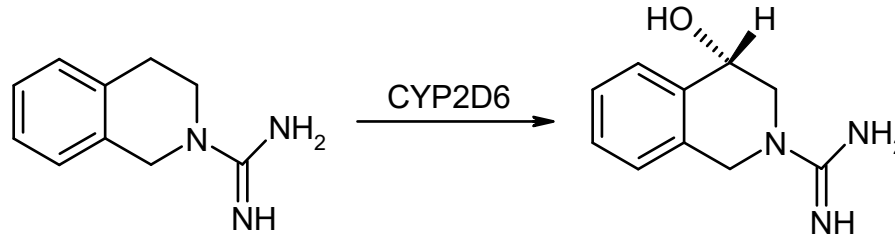
FAFLVSPSPYELCAVPR

T impaired metabolism of sparteine in alleles 2, 10, 12, 14 and 17 of CYP2D6

see <http://www.expasy.org/cgi-bin/niceprot.pl?P10635>

CYP 2D6 Polymorphism (IV)

variability of debrisoquine-4-hydroxylation



Lit: T. Winkler *Deutsche Apothekerzeitung* **140** (2000) 38

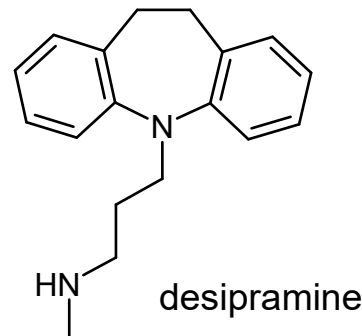
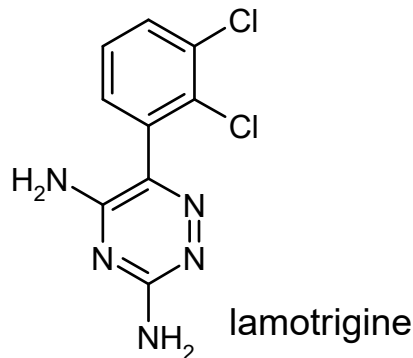
CYP 2D6 Polymorphism (V)

the *poor metabolizer* phenotyp has consequences for the metabolism of more than 25% of all common drugs, since it causes an increased concentration of xenobiotics that are not metabolized.

Lit: H.K.Kroemer & M.Eichelbaum. *Life Sci.* **56** (1995) 2285.

Thus, CYP2D6 genotyping is already applied to select appropriate test candiates in phase II of clinical tests:

lamotrigine, desipramine (Antidepressants)



Lit: M.P.Murphy et al. *Pharmacogenetics* **10** (2000) 583.

Polymorphism of further CYPs

CYP 1A1 ca. 16 isoforms (lung)

CYP 1A2 individual; strong, medium, and slow conversion of caffeine

CYP 2B6 absent in 3-4 % of the caucasian population

CYP 2C9 deficit in 1-3 % of the caucasian population

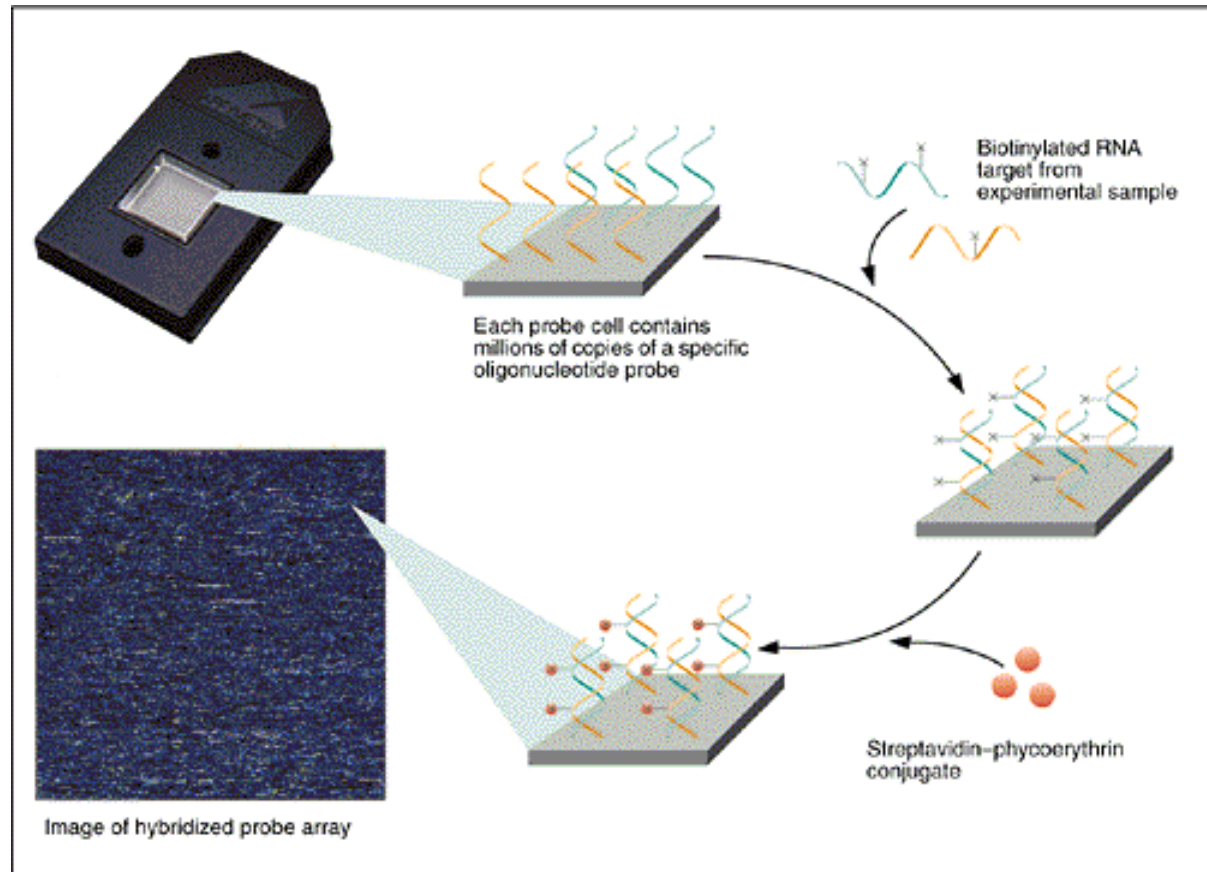
CYP 2C19 individuals with inactive enzyme (3-6 % of the caucasian and 15-20 % of the asian population)

CYP 2D6 poor metabolizers in 5-8 % of the european, 10 % of the caucasian and <1% in the japanese population. Overexpression (gene duplication) in parts of the african and oriental population

CYP 3A4 ca. 56 isoforms (liver)

Genotyping for P450 alleles

Affymetrix (US) has developed microarrays (gene chips) using immobilized synthetic copies of P450 nucleotides, that allow the identification of all clinically relevant allelic variants.



Induction and regulation of CYP3A (I)

A series of xenobiotics have been identified that lead to increased expression of enzymes of the CYP3A family.

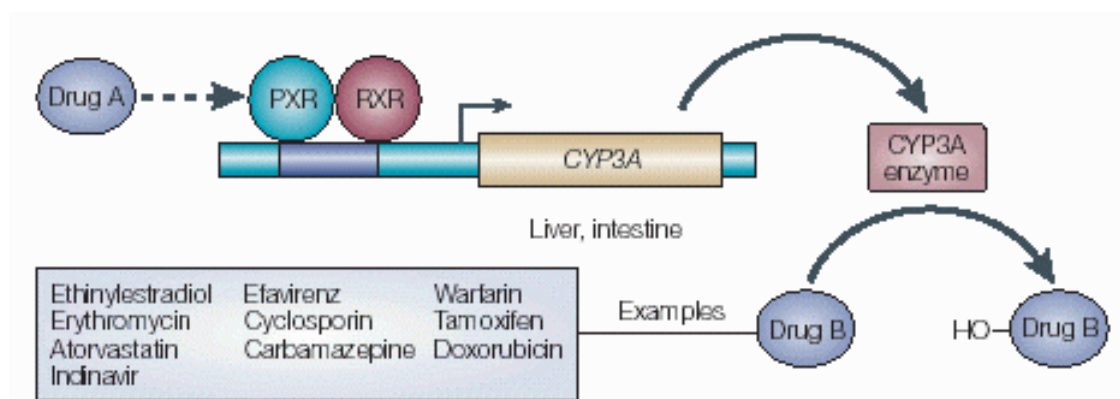
indinavir	antiviral
efavirenz	antiviral
cyclosporine	immuno-suppressant
carbamazepine	antipsychotic
atorvastatin	HMG CoA reductase inhibitor
tamoxifen	anti-hormone

These bind to the *pregnane X receptor* (PXR) which is the transcription factor for the regulation of the CYP3A gene expression.

Lit: T.M. Wilson et al. *Nature Rev. Drug Disc.* **1** (2002) 259

Induction and regulation of CYP3A (II)

The PXR receptor operates together with the *retinoid X receptor* (RXR) as a heterodimer.



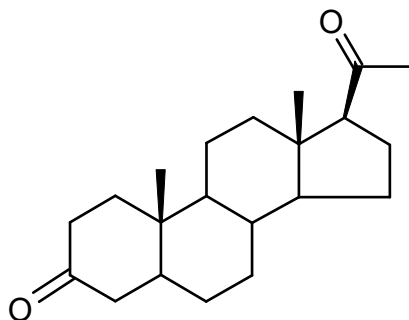
CYP3A induction leads to an increased metabolism of the administered substance due to upregulated enzymes. This can cause adverse reactions, such as inflammation of the liver (hepatitis).

Lit: T.M. Wilson et al. *Nature Rev. Drug Disc.* **1** (2002) 259

RXR and other nuclear receptors (I)

As a specific, endogen activator of RXR, 5 β -pregnane-3,20-dione has been identified.

In contrast, PXR is much less specific and is activated by glucocorticoids as well as by anti-glucocorticoids.



5 β -pregane-3,20-dione

Conversely, the unspecific *constitutive androgen receptor* (CAR) is found in the cytoplasm and dimerizes with PXR in the nucleus. Analog to PXR, the CYP2B gene is regulated.

Likewise high sequence homology has been found for the *vitamine D receptor* (VDR) that regulates CYP27, and for the *arylhydrocarbon receptor* (AHR) (dioxin receptor).

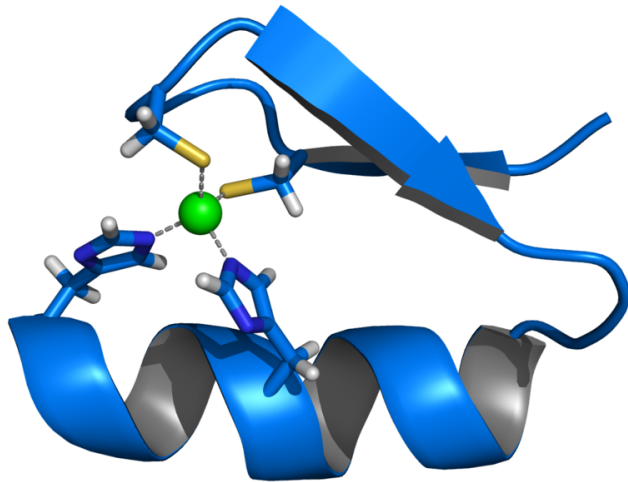
RXR and other nuclear receptors (II)

These nuclear receptors all belong to a family of transcription factors. Each one possess a double zinc-finger DNA-binding domain (DBD), and a larger ligand binding domain (LBD) which is located at the C-terminus.



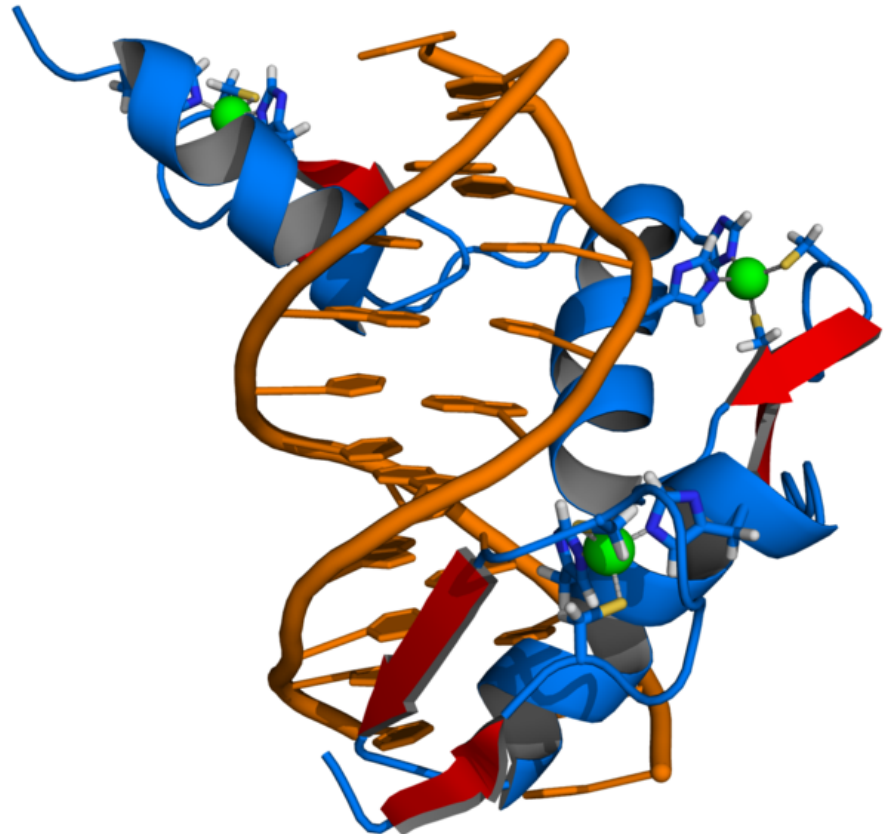
They have been initially called *orphan nuclear receptors* as their ligands have been found later.

Zinc finger motif in DNA-binding motifs



The zinc ion is coordinated by two cysteines and two histidines.

Source: Wikipedia



The protein Zif268 contains three zinc fingers motifs in complex with the DNA

Human Orphan Nuclear Receptors

receptor (gene ID)	natural ligand / synthetic ligand
CAR (NR1I3)	3 α ,5 α -androstanol
COUP (NR2F1)	β -naphthoflavone
ERR (NR3B1)	isoflavones, 4-hydroxytamoxifen (=afimoxifene)
FXR (NR1H4)	chenodeoxycholic acid (a bile acid), tropifexor
HNF4 (NR2A1)	palmitic acid, lauric acid
LRH (NR5A2)	phospholipids, GSK8470
PPAR (NR1C1)	eicosapentaenoic acid, hypolipidemic drugs, e.g. saroglitazar
PXR (NR1I2)	5 β -pregnane-3,20-dione, rifampicin
ROR (NR1F1)	all- <i>trans</i> retinoic acid, melatonin
RXR (NR2B1)	9- <i>cis</i> -RNA, adapalene, etodolac

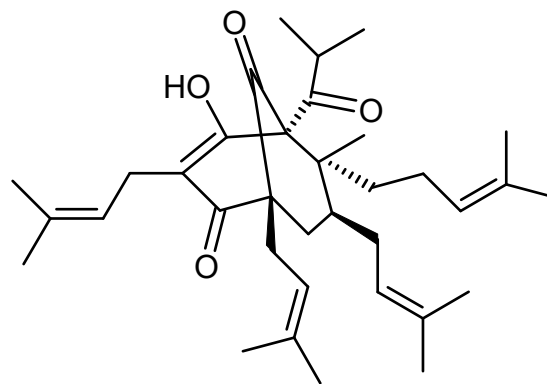
Selection only, for more see references:

Lit: T.M.Wilson & J.T.Moore *Mol. Endocrin.* **16** (2002) 1135.

I.M.S.de Vera *ACS Pharmacol. Transl. Sci.* **1** (2018) 134.

Additions from corresponding UniProt entries.

Induction and regulation of CYP3A (III)



hyperforin, a natural ingredient of St. John's wort (Johanniskraut, *Hypericum perforatum*) exhibits the highest measured affinity to PXR ($K_d = 27$ nM) so far.

Application: remedy against cholestasis [Gallestauung], mild antidepressant (heavily debated if available concentration in preparations of St. John's wort is sufficiently high)

Induction and regulation of CYP3A (IV)

X-ray structure of PXR complexed with hyperforin (1M13.pdb)



Lit: R.E. Watkins et al. *Biochemistry* **42** (2003) 1430

Induction of further CYPs

CYP 1A2 omeprazole, insulin, aromatic hydrocarbons
(cigarette smoking, charbroiled meat)

causes increased caffeine level in the plasma,
if you quit smoking.

CYP 2C9 rifampicin, secobarbital

CYP 2C19 carbamazepine, prednisone

CYP 2D6 dexamethason

CYP 2E1 ethanol, isoniazid

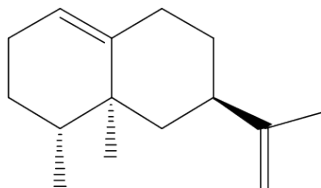
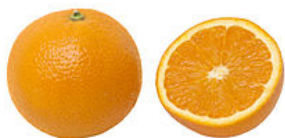
CYP 3A4 glucocorticoides, phenobarbitone,
rifampicin, nevirapine, sulfadimidine,
nevirapine, sulfinpyrazone, troglitazone

Typical inhibitors of various CYPs

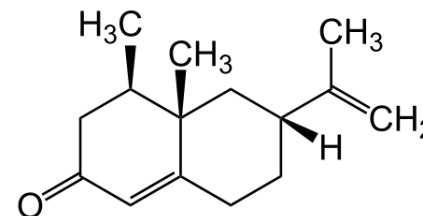
CYP 1A2	cimetidine, ciprofloxacin, enoxacin... grapefruit juice (naringin, 6',7'-dihydroxy-bergamottin)
CYP 2C9	chloramphenicol, amiodarone, omeprazole,...
CYP 2C19	fluoxetine, fluvastatin, sertraline,...
CYP 2D6	fluoxetine, paroxetine, quinidine, haloperidol, ritonavir,...
CYP 2E1	disulfiram, cimetidine,...
CYP 3A4	cannabinoids, erythromycin, ritonavir, ketoconazole, grapefruit juice

Biotechnical Application of CYPs

Synthetic steps that cannot be carried out by conventional chemistry



valencene
from orange peels



(+)-nootkatone
0.5 vol.% in grapefruit oil; 1g: 118€

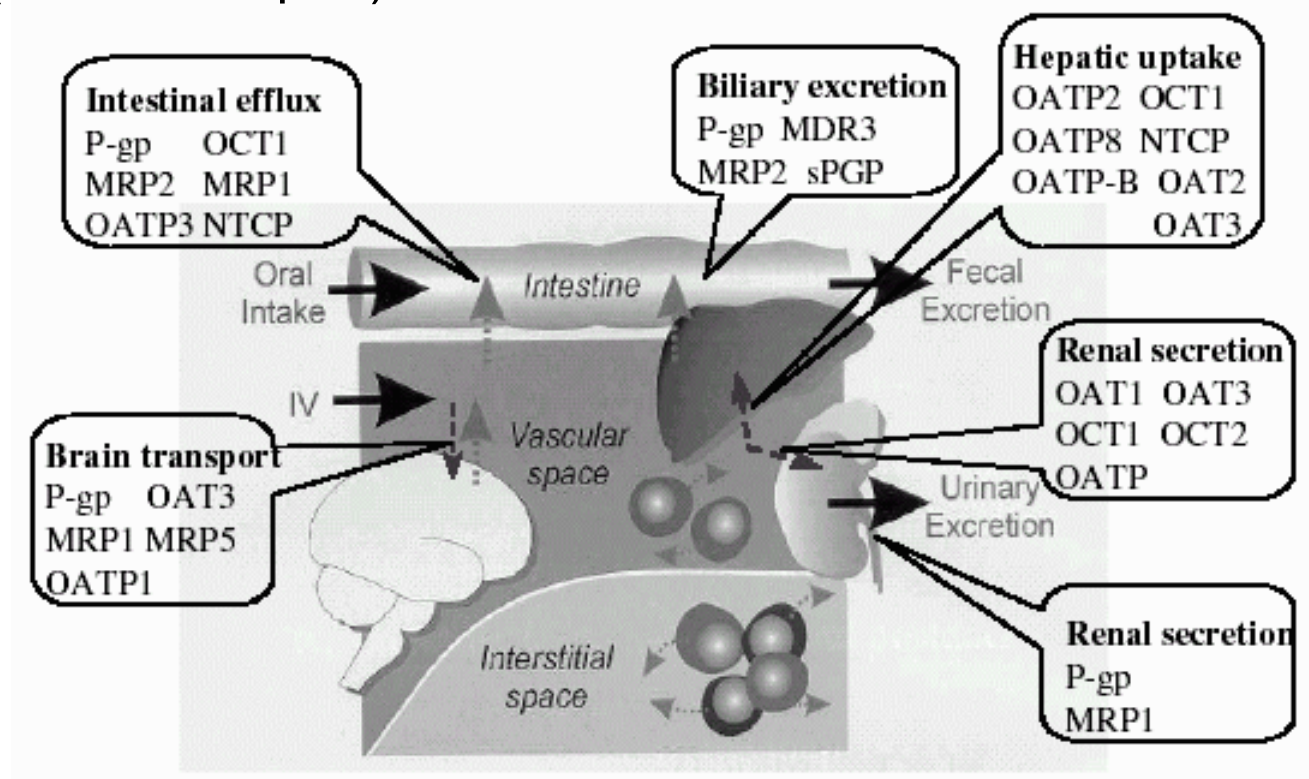
Synthesis of expensive aroma compounds (nootkatone, limonene, citral, ionone) from cheaper natural products.

Synthesis of drug metabolites as test standards.

Picture sources: wikipedia.org

Transporters (I)

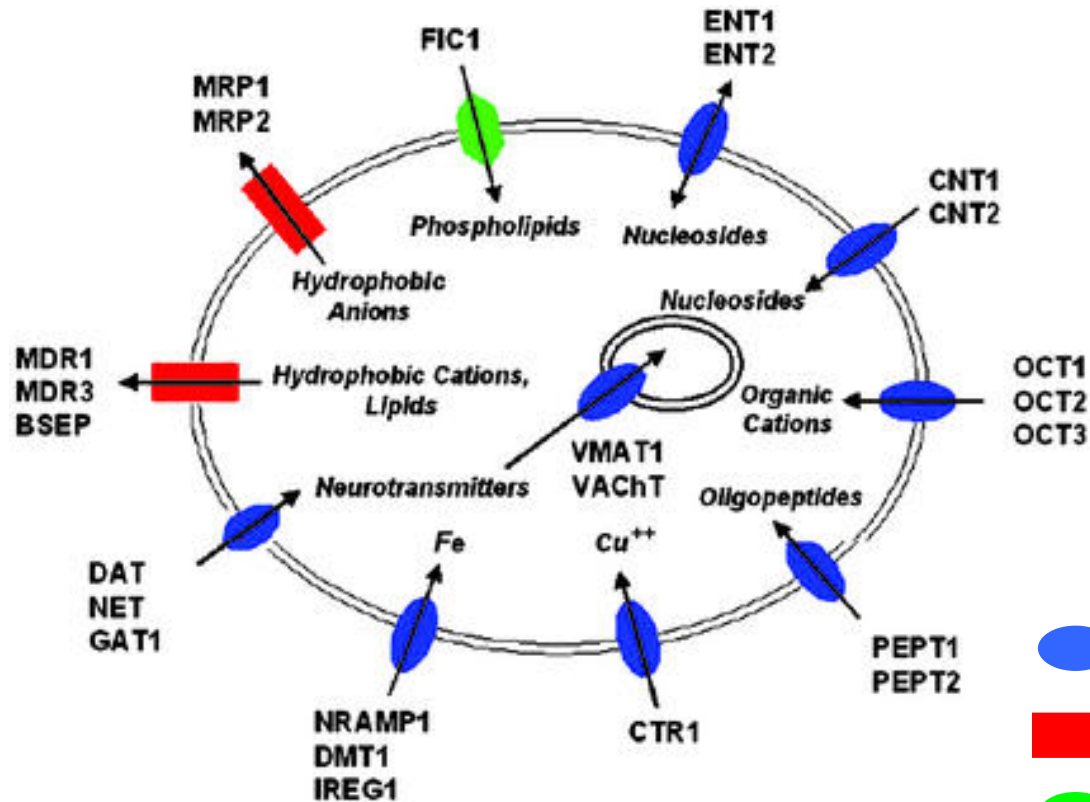
In contrast to the passive diffusion through membranes transporters cause increased *influx* into, or conversely *efflux* from compartments, whereby ATP is consumed. (active transport)



Lit: A.Ayrton et al. *Xenobiotica* **31** (2001) 469

Transporters (II)

Membrane bound transporters involved in the pharmacokinetic of endogenous substances



superfamilies:

● solute carriers (SLC)

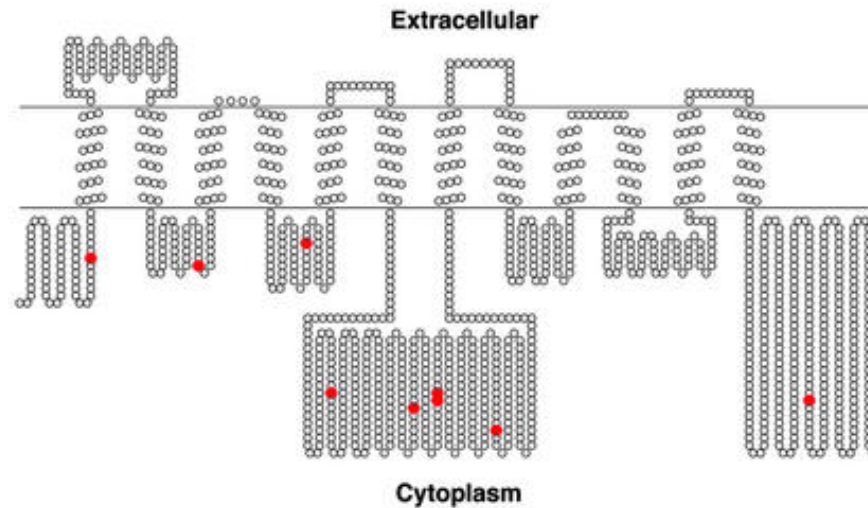
■ ATP-binding cassette (ABC)

● P-type ATPase

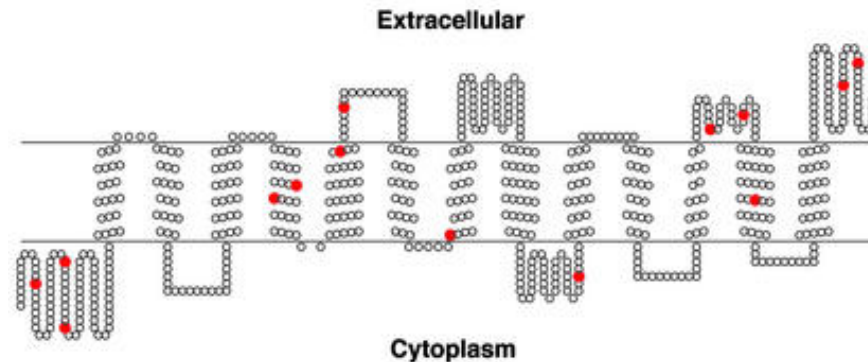
Lit: M.K.Leabman et al. *Proc.Nat.Acad.Sci.USA* **100** (2003) 5896

Structure of membrane-bound transporters

BILE SALT EXPORT PUMP (ABCB11)



CONCENTRATIVE NUCLEOSIDE TRANSPORTER 1 (SLC28A1)



Membrane-bound transporters are proteins with up to 12 and more transmembrane helices that are connected by loops. So far no X-ray structure of a transporter has been achieved.

Lit: M.K.Leabman et al. *Proc.Nat.Acad.Sci.USA* **100** (2003) 5896

P-glycoprotein (P-gp pump, MDR1)

P-gp belongs to the group of *multidrug resistant proteins* (MDR) and is encoded by the ABCB1 gene.

It is an ATP-dependent efflux pump and transmembrane protein.

Especially the bioavailability of antipsychotics is limited by the mediated efflux from the brain and central nervous system back into the system blood circulation.

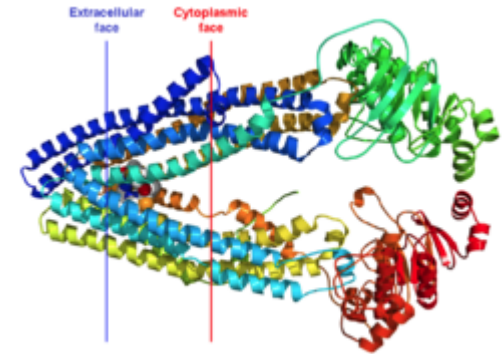
Likewise transport of substances from the liver into the gastrointestinal (*biliary excretion*) e.g. of indinavir

Overexpression of P-gp in cancer cells leads to resistance against antineoplastics.

Lit: A.Ayrton et al. *Xenobiotica* **31** (2001) 469.

A.Seelig *Eur.J.Biochem.* **251** (1998) 252.

picture source: wikipedia.org



Transporter proteins for organic ions

Comprising the families of the

Organic Anion Transporters (OAT) and the

Organic Cation Transporters (hOCT1-3, hOCTN1-2)

They contribute in particular to the hepatic uptake and renal excretion of hydrophilic metabolites and catabolites.

Population-specific SNPs may lead to strong differences regarding dosage of certain drugs (e.g. rosuvastatin: OATP1B1 organic anion transporting polypeptide 1B1)

Lit: A.Ayrton et al. *Xenobiotica* **31** (2001) 469.

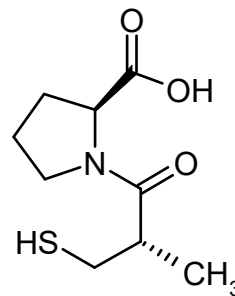
H.Koepsell *Pharmacol. Rev.* **72** (2020) 253.

Transporter proteins for *influx*

There are also transporters that mediate the active uptake of substance from the intestine

PepT1 (intestinal peptide transporter 1, SLC15A1)
transmembrane protein possessing 12 TM-helices
Responsible for the uptake of nitrogen!

substrates: small peptides (di- and tripeptide, as well as compounds exhibiting peptide-like features, e.g. captopril)



Polymorphisms of transporters

Also transporters show considerable genetic variations:

gene	protein / function
------	--------------------

ABCB1	(ATP-binding cassette subfamily B member 1) P-gp efflux
-------	---

SLC6A3	(dopamine transporter) neurotransmitter
--------	---

SLC6A4	(serotonin transporter) neurotransmitter
--------	--

ADRB2	(β -adrenergic receptor) receptor for β -blockers
-------	--

ALOX5	(arachidonate 5-lipoxygenase) biosynthesis of leukotrienes
-------	---

See: D.B.Goldstein et al. *Nature Rev. Genetics* **4** (2003) 937.