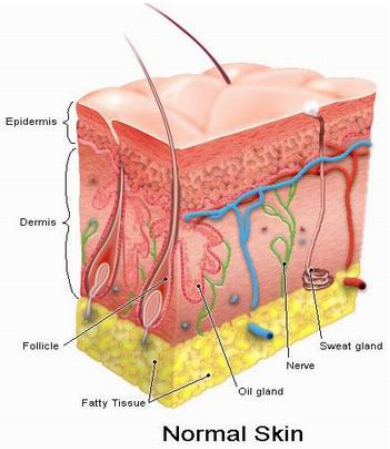
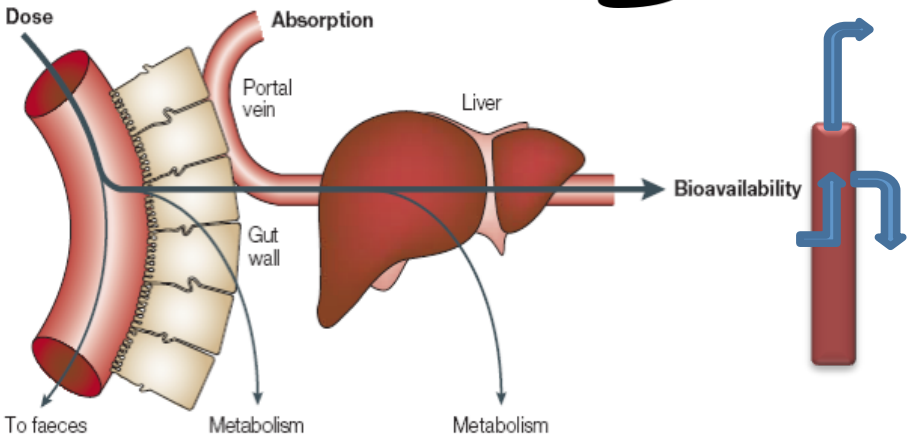
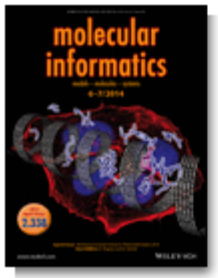
Investigating and Predicting how Biology Changes Molecules and their Properties







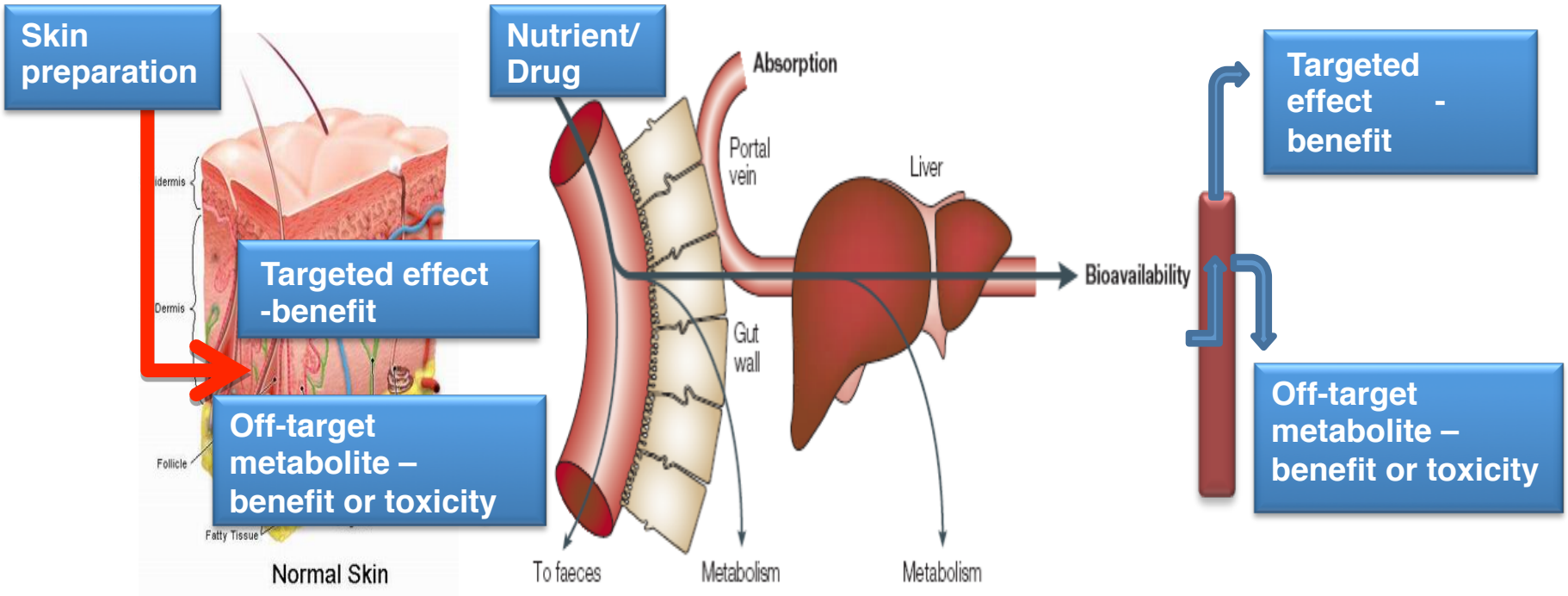
Molecular Informatics
Special Issue: Strasbourg Summer School in Chemoinformatics 2014
Volume 33, Issue 6-7, pages 443–445, June 2014
DOI: 10.1002/minf.201400031

Robert Glen

Imperial College London
UNIVERSITY OF CAMBRIDGE

Metabolism and Transport: key biological properties of medicines

An example of the journey of a functional molecule – orally, or from skin – we must optimise the pharmacokinetics, safety and efficacy – metabolism and transport are key factors to take into account

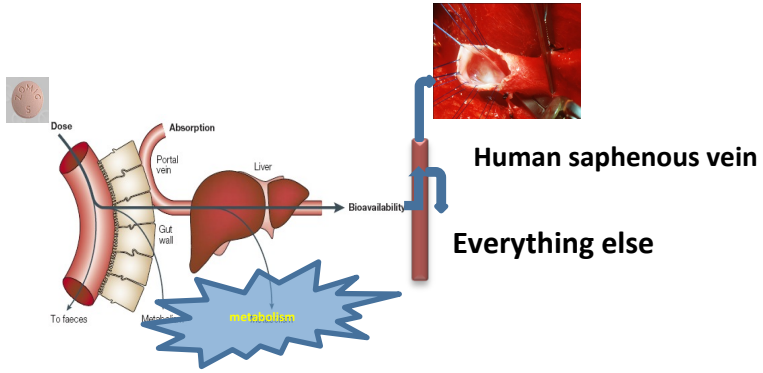


Predicting or modelling metabolism and transport would give significant gains in drug discovery

Metabolism and Transport

- Metabolism has many key roles including
  - Anabolism (making molecules)
  - Catabolism (breaking molecules down)
  - Detoxification of toxic molecules
  - Elimination of molecules from the system
  - Changing bioactivity – e.g. to invoke signalling
- Transport
  - Special transport Proteins recognise specific molecules
  - Transporters are expressed only in specific cells
  - Transport mediated processes dominate the movement of molecules, including xenobiotics
  - Gatekeepers of homeostasis

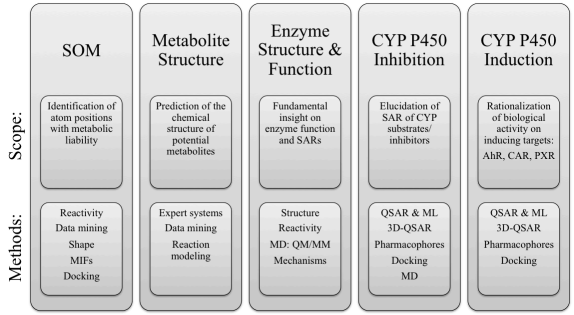
Metabolism



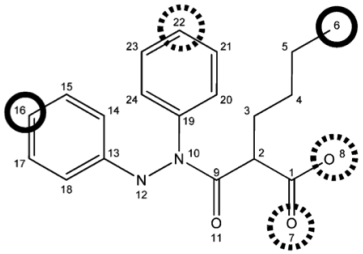
- Understanding pharmacokinetics/toxicity of compounds (functional actives, drugs etc.) is very important.
  - ADMET – PK - about 10% of failures of potential drugs (getting lower)
    Toxicity – about 40% of failures (getting higher as risk/benefit needs to be almost infinity! (zero/one).
- Many xenobiotics and toxins are highly lipophilic
  - Typically are non-polar to cross biological membranes
  - However, must be polar to be excreted (kidneys)
  - Metabolic steps often reduce lipophilicity of molecules
- Metabolism can:
  - Alter activity e.g. antagonist to agonist
  - Deactivate/activate molecules
  - Convert pro-drugs/substrates into active forms
  - Produce toxic compounds/induce DDI's (drug/drug interactions)
  - Create environmental toxins – e.g. Endocrine disruptors

Prediction of Metabolism

- Therefore, there is a longstanding interest in predicting the metabolic fate of molecules using both simulation and informatics approaches
- We have developed :
  - database of enzyme mechanisms (Macie) available at EBI
  - software for prediction of products of metabolism (MetaPrint2D), FAME,
  - and recently mechanism-based methods for prediction of sites of metabolism in Cytochrome-P450 enzymes.



Computational Prediction of Metabolism: Sites, Products, SAR, P450 Enzyme Dynamics, and Mechanisms. Kirchmair et al. J. Chem. Inf. Model., 2012, 52 (3), pp 617–648 DOI: 10.1021/ci200542m



Anti-cancer Drug Development: Computational Strategies to Identify and Target Proteins Involved in Cancer Metabolism. Mak, L. Et al. Current Pharmaceutical Design, Vol19(4), 2013 , 532-577.

Databases of Metabolism

- Databases of metabolism are diverse – from maps of metabolic pathways to likely metabolic products to metabonomics of body fluids. There is much to learn (and derive models) from curated databases of metabolic processes.
- Some examples:

The NMR metabolomics database of Linkoping, Sweden (MDL). An on-line database and publically accessible depository dedicated to the omics of small biomolecules. It is intended to facilitate access to NMR parameters of small metabolites in liquid phase (aqueous solutions only).

Biological Magnetic Resonance Bank (BMRB). This Metabolomics database is available to the NMR community.

The Human Metabolome Database (HMDB). An electronic database containing detailed information about small molecule metabolites found in the human body.

PharmGKB, Encyclopaedic database focussed on drug metabolites

The Madison Metabolomics Consortium Database (MMCD). This database, maintained by the National Magnetic Resonance Facility at Madison, is a resource for metabolomics research based on NMR spectroscopy and mass spectrometry.

The Brüschweiler Laboratory COLMAR Metabolomics Web Portal . Complex Mixture Analysis by NMR (COLMAR)

KEGG pathway database (*KEGG*). A collection of manually drawn pathway maps representing our knowledge on the molecular interaction and reaction networks . Most widely used.

BioCyc. A collection of 371 Pathway/Genome Databases

MetaCyc. Metabolic pathways and enzymes from more than 900 organisms

GOSTAR. Metabolites of drugs from 50,000 publications

GeneMedRx http://www.genemedrx.com/drug-metabolism.php Drug/Drug interactions

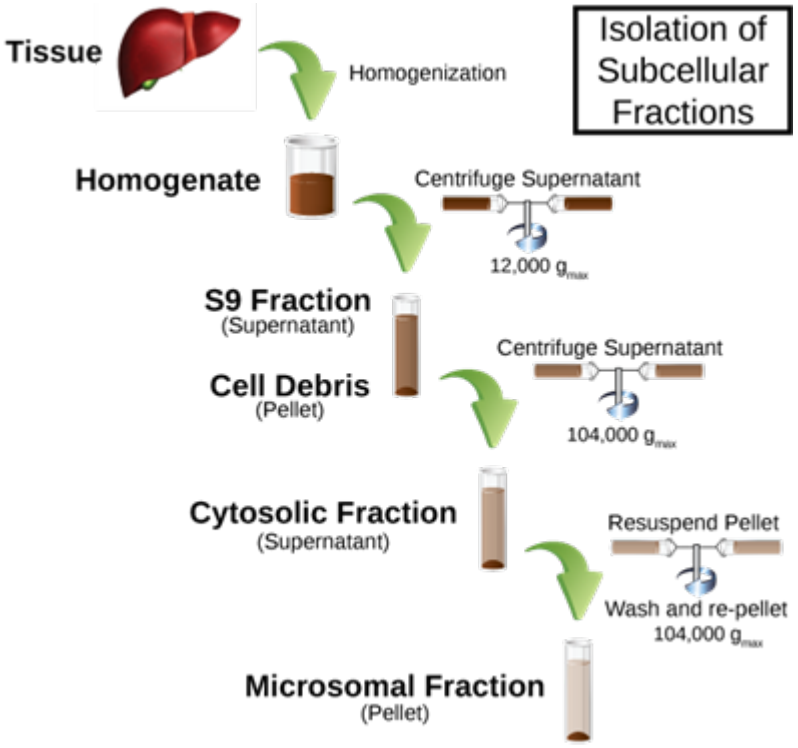
Metabolite http://accelrys.com/products/databases/bioactivity/metabolite.html substrates and products . 100,000 transformations

-----etc-----

MACIE http://www.ebi.ac.uk/thornton-srv/databases/MACiE/

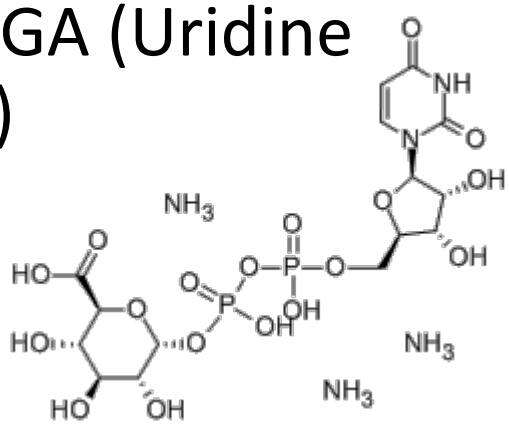
Where does the data come from?

- Subcellular systems
  - Organelles
  - Homogenate fractions - S9 (post-mitochondrial supernatant fraction) consists of microsomes and cytosol.
  - Blood serum and plasma
- Cellular and tissue systems
  - Primary cell cultures - hepatocytes
  - Tumor cell lines – Caco-2 cells
  - Tissue slices
  - Isolated perfused organs
- *In-vivo* systems
  - Multicellular organisms
  - Batches of experimental animals
  - Groups of individuals
  - Collectives of patients
  - Populations



Data on metabolic stability and sites of metabolism

- Incubations with individual CYP P450s
  - DD interactions, mechanism-based inhibition
- Hepatic Microsomal incubations
  - Oxidative metabolism
- Hepatic microsomal incubations + UDPGA (Uridine 5'-diphospho-glucuronosyltransferase)
  - Conjugation reactions (Phase-II)



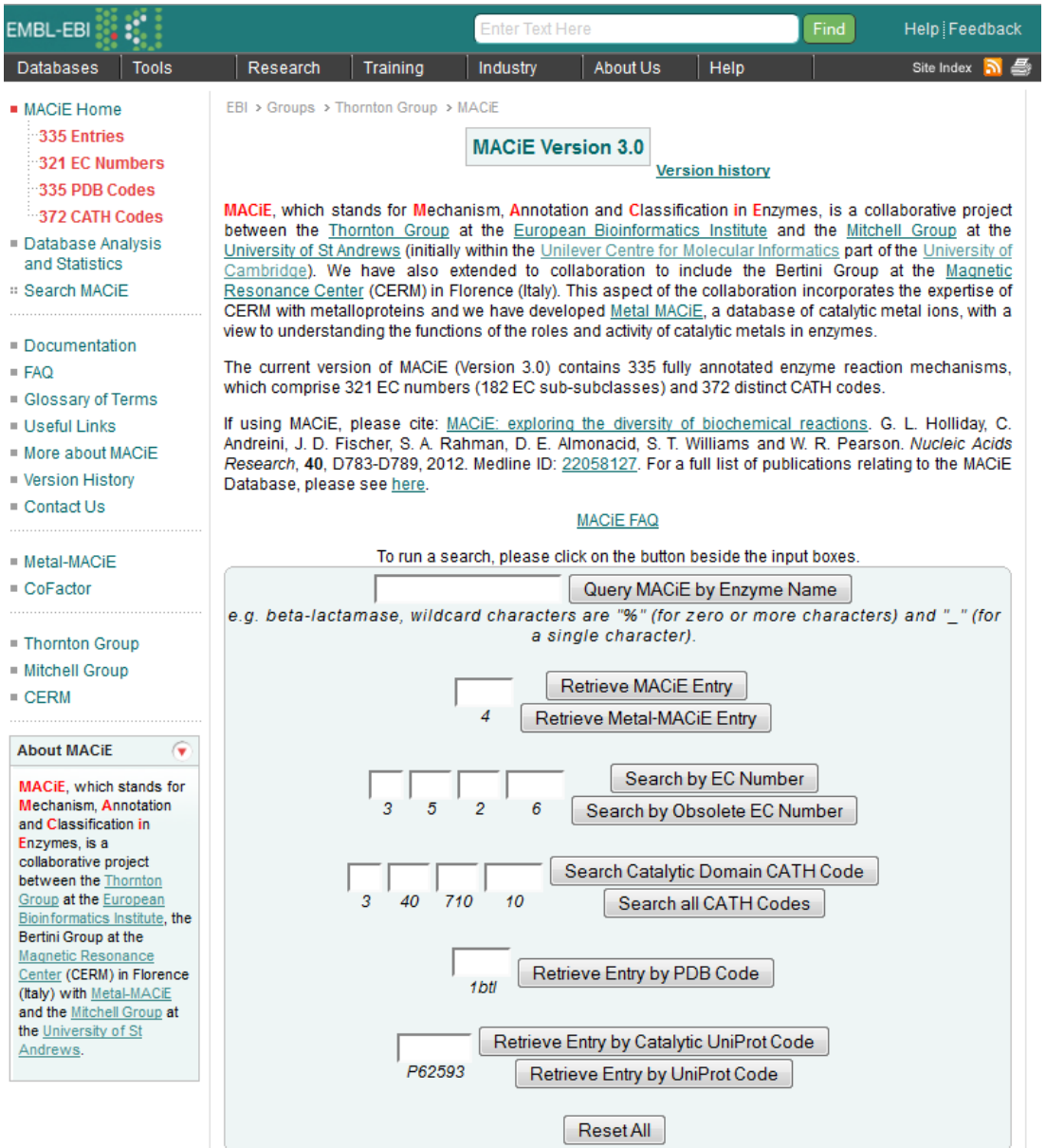
- Reactive metabolite trapping
  - Trapped using glutathione or cysteine
- Animal models
  - Mouse/rat may be humanised. Metabolism/transport and everything else!

Characterisation and detection

- Radio labelled species may be used with scintillation counting – sensitive, specific, need to be synthesised
- LC-MS is commonly used for metabolite identification/ quantification with standards
- LC-NMR/NMR-MS used for structure determination

- Note – this is often difficult and time consuming, not all metabolites may be detected (and characterised) or reported in papers. Experimental methods differ, so detection can vary.

MACIE : Mechanism, Annotation and Classification in Enzymes



Mechanism, Annotation and Classification in Enzymes is a collaborative project between the Thornton Group at the European Bioinformatics Institute and the Mitchell Group at the University of St Andrews (previously Unilever Centre for Molecular Informatics part of the University of Cambridge).

Gemma Holiday and John Mitchell

MACiE: exploring the diversity of biochemical reactions. G. L. Holliday et al., Nucleic Acids Research, 40, D783-D789, 2012.

Query for entry M0210 3.5.2.6 beta-lactamase (Class D)

Overview for MACiE Entry M0210

Version history

General Information

EC Number: 3.5.2.6 (*A member of the Hydrolases, Acting on carbon-nitrogen bonds, other than peptide bonds, In cyclic amides*)

Enzyme Name: beta-lactamase (Class D)

Biological Species: *Escherichia coli* (Bacteria)

Catalytic Chain UniprotKB Accession Codes:

- P13661 - *Beta-lactamase OXA-1*

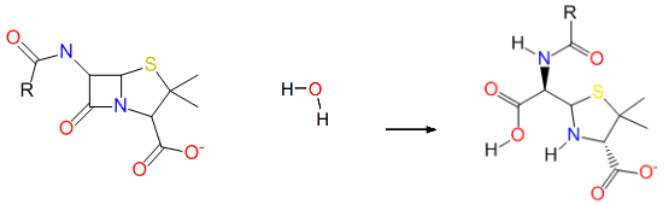
Representative PDB Code: 1m6k - *STRUCTURE OF THE OXA-1 CLASS D BETA-LACTAMASE* (Resolution = 1.50 Å).

Catalytic CATH Codes:

- 3.40.710.10 - *DD-peptidase/beta-lactamase superfamily*

Display structure information

Overall Reaction:



beta-Lactam	water	substituted beta-amino acid
C01866	C00001	C03806
CHEBI:35627	CHEBI:15377	CHEBI:33705

Overall Comment: Kcx70 is a lysine residue that has been carbamylated, this has been shown to be the general acid/base in the Class D beta-lactamases and compensates for the lack of a general base to correspond to the Glu166 in Class A beta-lactamases [1,2].

View similar reactions

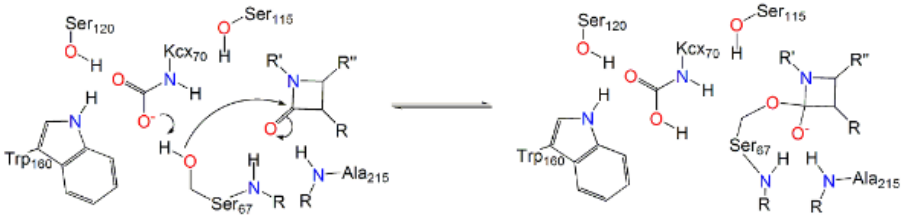
Stepwise Description of the Reaction

Step 1	Carbamylated lysine (Kcx70) deprotonates the alcohol of Ser67, which initiates a nucleophilic addition to the carbonyl group of the betalactam ring.
Step 2	The tetrahedral intermediate collapses, cleaving the C-N bond, which deprotonates Ser115, which in turn deprotonates the secondary amine of Kcx70, which deprotonates the carboxyl group of Kcx70.
Step 3	Kcx70 deprotonates water, which attacks the carbonyl carbon of the covalently attached intermediate in a nucleophilic addition.
Step 4	The tetrahedral intermediate collapses, eliminating Ser67, which reprotonates from Kcx70, producing the product.

Query for entry M0210 3.5.2.6 beta-lactamase (Class D)

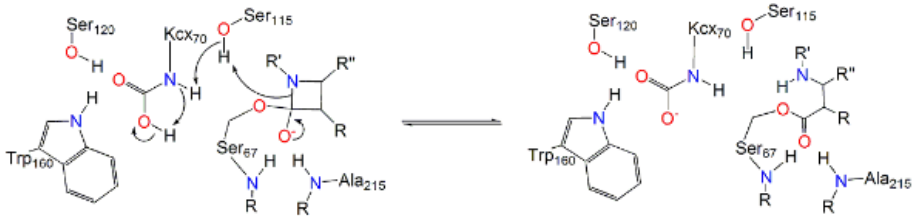
Step 01

Carbamylated lysine (Kcx70) deprotonates the alcohol of Ser67, which initiates a nucleophilic addition to the carbonyl group of the betalactam ring.



Step 02

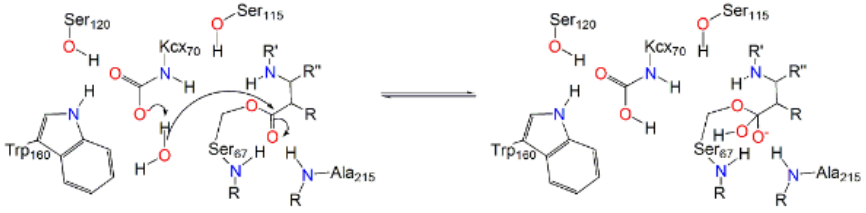
The tetrahedral intermediate collapses, cleaving the C-N bond, which deprotonates Ser115, which in turn deprotonates the secondary amine of Kcx70, which deprotonates the carboxyl group of Kcx70.



Comment: A conserved serine (Ser115) may assist in transferring the proton to the nitrogen of the thiazole leaving group. However, given the 4-5 Angstrom distance between Ser115 and the carboxylate group, it is probably necessary for the NE of Lys70 to participate in the proton transfer [1].

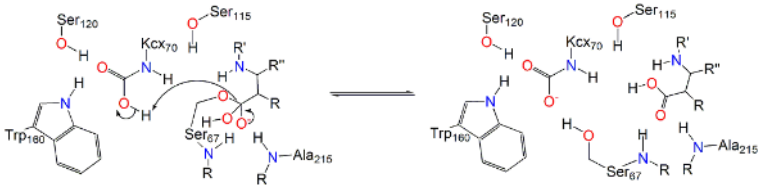
Step 03

Kcx70 deprotonates water, which attacks the carbonyl carbon of the covalently attached intermediate in a nucleophilic addition.



Step 04

The tetrahedral intermediate collapses, eliminating Ser67, which reprotonates from Kcx70, producing the product.



Mechanisms

Proton Transfer
Unimolecular Elimination by the Conjugate Base

Mechanism Components

Bond Cleavage
Bond Formation
Bond Order Change
Enzyme-Substrate Bond Cleavage
Intermediate Collapse
Intermediate Terminated
Overall Product Formed
Enzyme Regenerated

Amino acids involved in the reaction step.

Amino Acid	Location of Function	Activity	Function
Ser115A	Side Chain	spectator	Hydrogen Bond Donor Hydrogen Bond Acceptor
Ala215A	Main Chain Amide	spectator	Hydrogen Bond Donor Electrostatic Stabiliser
Trp160A	Side Chain	spectator	Hydrogen Bond Donor Electrostatic Stabiliser
Ser67A	Main Chain Amide	spectator	Hydrogen Bond Donor Electrostatic Stabiliser
Kcx70A	Side Chain	reactant	Hydrogen Bond Acceptor Hydrogen Bond Donor Proton Donor
Ser67A	Side Chain	reactant	Covalently Attached Hydrogen Bond Acceptor Proton Acceptor Nucleofuge
Ser120A	Side Chain	spectator	Hydrogen Bond Donor Electrostatic Stabiliser

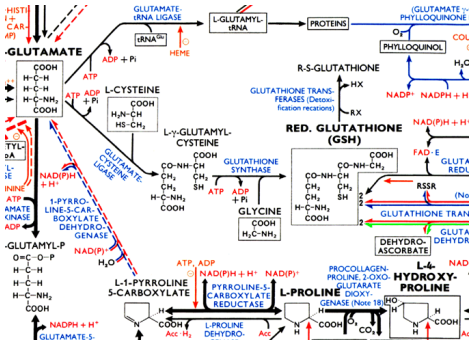
Reactive Centre

Bonds Formed	Bonds Cleaved	Bonds Changed in Order	Atom Types Involved
O-H	C-O O-H	The C-O bond changes from a single to double bond	C H O

Predicting metabolism

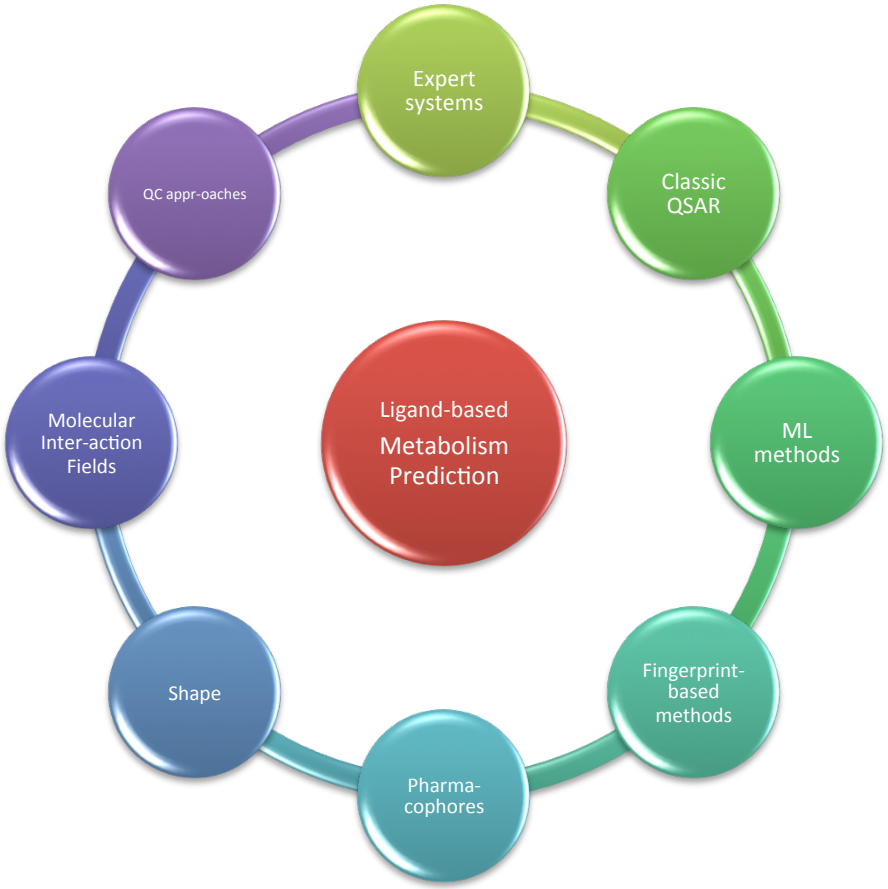
Metabolism is often classified as : Phase-1 (typically oxidation) and Phase-2 (conjugation) reactions

- Phase I metabolism – e.g.
  - Oxidative Reactions:
    - Aromatic hydroxylations, alkene epoxidations, C adjacent to sp2 centres, aliphatic or alicyclic C oxidations, C-N oxidations, O-dealkylation, C-S oxidations, et al (dehalogenation, aromatization, oxidation of arenol)
  - Reductive Reactions:
    - Carbonyl reductions, nitro reduction, azo reduction, tertiary amine oxide, dehalogenation
  - Hydrolytic Reactions:
    - acid or base hydrolysis of esters and amines giving carboxylic acids, alcohols & amines
- Phase II metabolism – e.g.
  - Glucoronidation, sulphation, glycine conjugation and others.



Silverman, R.B. (1992) The Organic Chemistry of Drug Design and Drug Action. Academic Press Inc., San Diego USA.

Prediction of metabolism typically falls into solving two problems

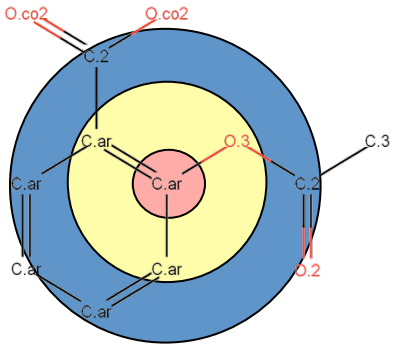


- Identify the site of metabolism (SOM)
- Identify the products
- Often seen as two separate problems.
- Determining flux in complex metabolic pathways is a separate problem

- We can take a
  - data-driven approach (read across, frequency analysis, rule-based)
  - SAR based
  - Phenomenological (simulation)

- Metaprint2D – a data driven approach

We wished to predict the sites and products of metabolism.
Initial approach based on molecular similarity of atom Environments (circular fingerprints)



Level			
0	**1**	**2**	**3**
C.ar	C.ar C.ar O.3	C.2 C.2 C.ar C.ar	C.3 C.ar O.2 O.co 2 O.co 2

SYBYL (Tripos Forcefield) atom types

Quite popular for many uses:

NMR - W. Bremser, HOSE—A novel substructure code, *Anal. Chim. Acta* ***103,*** 355–365 (1978).

pKa - Xing, Li; Glen, Robert C. Novel methods for the prediction of logP, pKa, and logD. J. Chem. Inf. Model. 2002, 42(4), 796-805.

Virtual screening - David Rogers and Mathew Hahn. Extended-Connectivity Fingerprints. J. Chem. Inf. Model., 2010, 50 (5), pp 742–754. Andreas Bender, Hamse Y. Mussa, and Robert C. Glen, Stephan Reiling. Similarity Searching of Chemical Databases Using Atom Environment Descriptors (MOLPRINT 2D): Evaluation of Performance. J. Chem. Inf. Comput. Sci., 2004, 44 (5), pp 1708–1718

Off target effects - Josef Scheiber et al. J. Gaining Insight into Off-Target Mediated Effects of Drug Candidates with a Comprehensive Systems Chemical Biology Analysis. Chem. Inf. Model. 2009, 49, 308–317

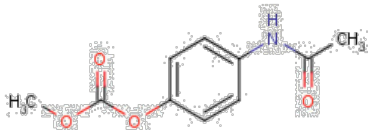
Metabolism – Scott Boyer, Catrin Hasselgren Arnby, Lars Carlsson, James Smith, Viktor Stein and Robert C. Glen. Reaction Site Mapping of Xenobiotic Biotransformations. J. Chem. Inf. Model. 2007, 47(2), 583-590.

Lars Carlsson , Ola Spjuth , Samuel Adams , Robert C Glen and Scott Boyer. Use of Historic Metabolic Biotransformation Data as a Means of Anticipating Metabolic Sites Using MetaPrint2D and Bioclipse. BMC Bioinformatics 2010, 11:362.

Metabolic Site/Product predictor (MetaPrint2D)

1 Symyx Metabolite database (~100,000 transformations) → Substrate + Products → Calculate environment for each substrate atom; Identify reaction centres

2 Query compound → Calculate environment for each atom

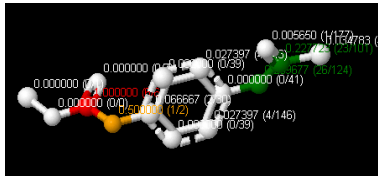


3 For each query atom, find all similar environments in database → How often is environment found at a reaction centre?

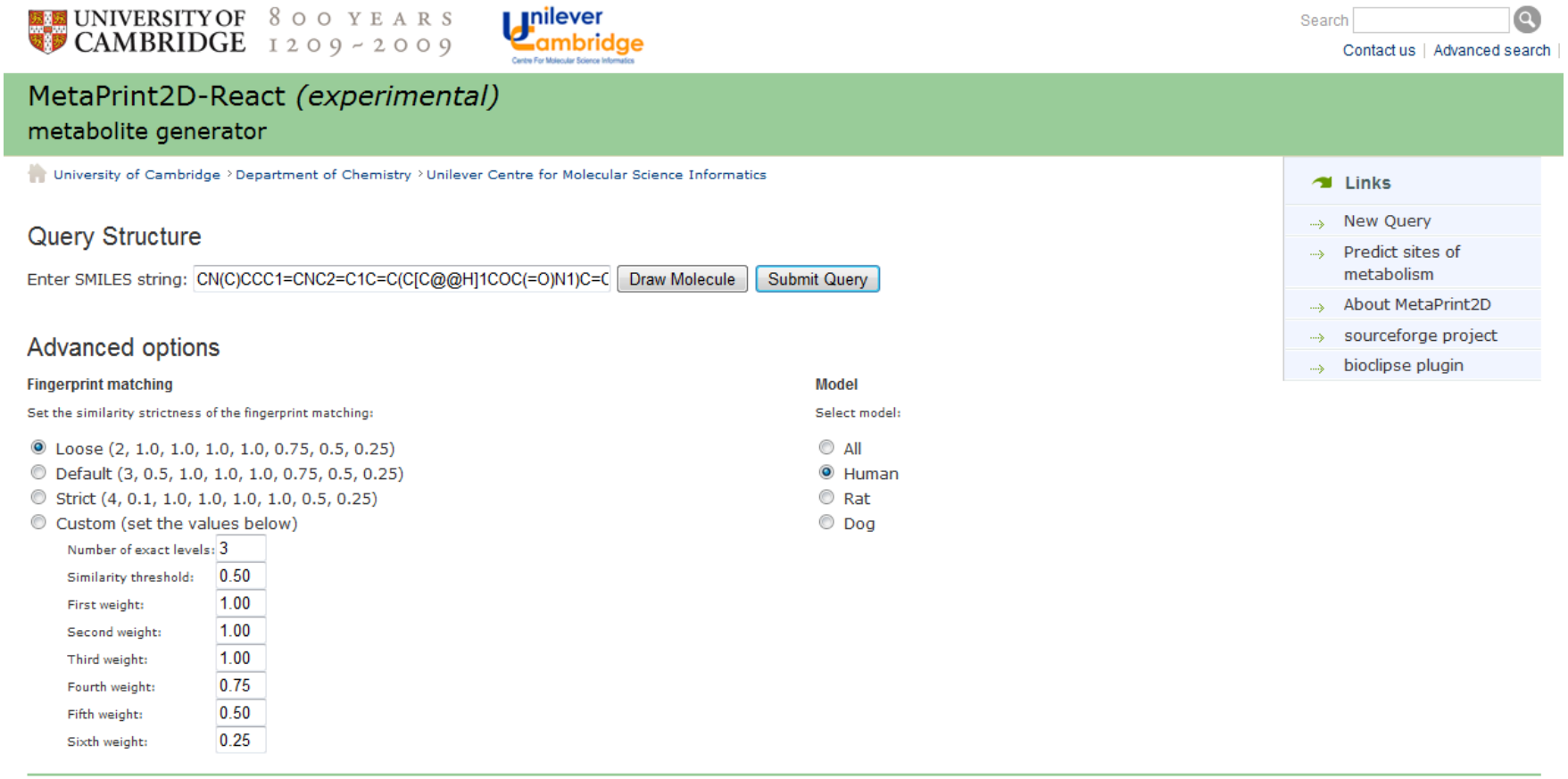
4 Calculate reaction occurrence ratios →

$$\frac{\text{Total number of similar reaction centres}}{\text{Total number similar atoms in rest of database}}$$

5 Calculate relative ratios for each atom in query compound, and display predictions Using a naive Bayes probabilistic model



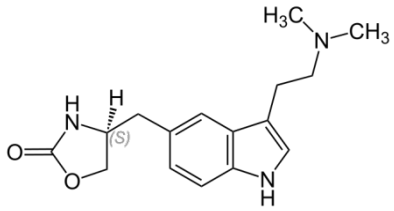
Metabolic Site/Product predictor (MetaPrint2D)



http://www-metaprint2d.ch.cam.ac.uk/metaprint2d/

This approach allows prediction of sites and probable products, including Phase II, in human, rat and dog. (not just cytochrome p450 metabolism)

Simple example of zolmitriptan, a migraine drug.



UNIVERSITY OF CAMBRIDGE 800 YEARS 1209–2009 | Unilever Cambridge

Search | Contact us | Advanced search

MetaPrint2D-React *(experimental)*
metabolite generator

University of Cambridge › Department of Chemistry › Unilever Centre for Molecular Science Informatics

Links
- New Query
- Predict sites of metabolism
- About MetaPrint2D
- sourceforge project
- bioclipse plugin

Instructions

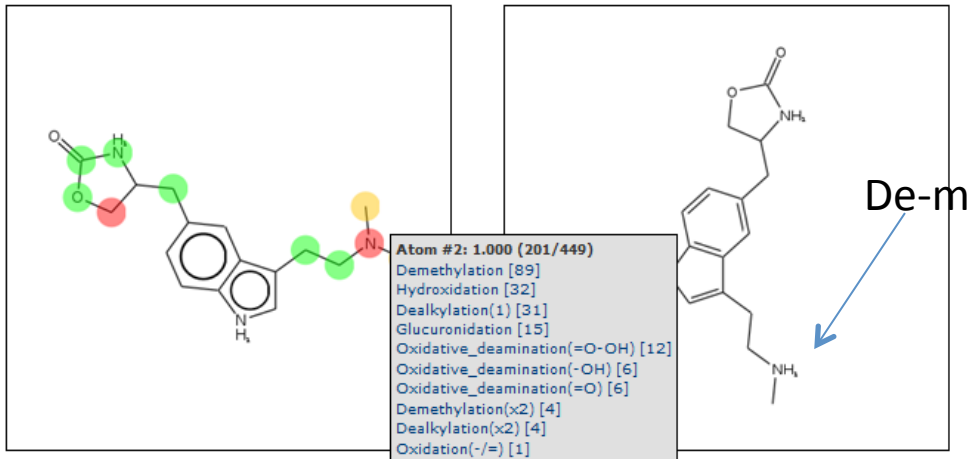
Click on an atom to display the predicted reactions at that site, and click on one of these reaction types to show the metabolite formed. Alternatively, select a reaction type from the filter list below the results image, to limit predictions to that reaction type.

The colour highlighting an atom indicates its normalised occurrence ratio (NOR). A high NOR indicates a more frequently reported site of metabolism in the metabolite database.

Red	0.66 <= NOR <= 1.00
Orange	0.33 <= NOR < 0.66
Green	0.15 <= NOR < 0.33
White	0.00 <= NOR < 0.15
Grey	Little/no data

Note: The normalised occurrence ratio does not indicate how likely a molecule is to be metabolised, but rather the relative likelihood of metabolism occurring at a particular site in the molecule, assuming it is metabolised.

Results



De-methylation

Reaction type filter: all

Input

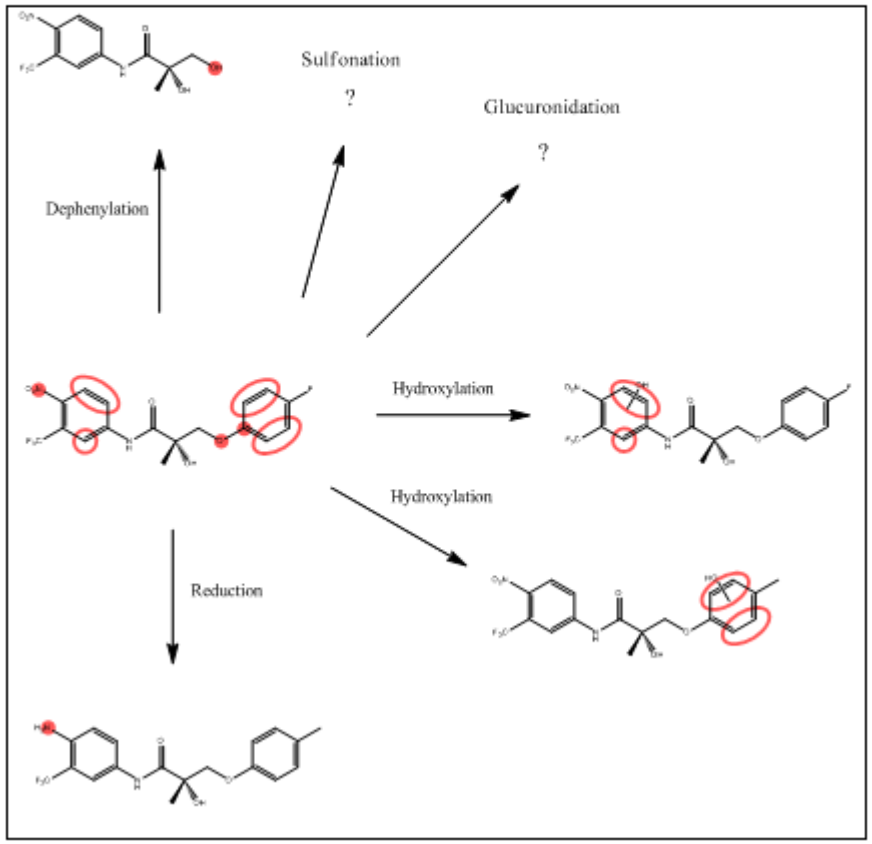
SMILES: CN(C)CCC1=CNC2=C1C=C(C[C@@H]1COC(=O)N1)C=C2
Model: METAB20081HUMAN
Settings: LOOSE

Interestingly, zolmitriptan (which has excellent bioavailability) is a partial agonist, while the main metabolite is a full agonist. So, as the drug concentration lowers in blood, the remaining compound becomes more potent – probably a longer lasting effect

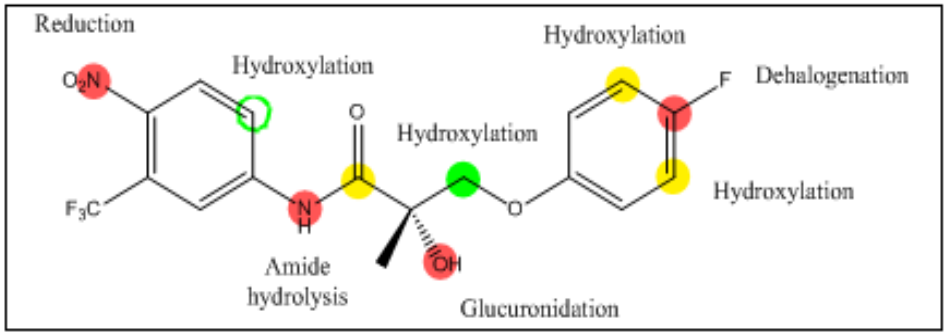
More complex example

6.10 Aryl-Propionamide-Derived Selective Androgen Receptor Modulator (*292*)

6.10.1 Reported metabolites



6.10.2 MetaPrint2D-React predicted transformations

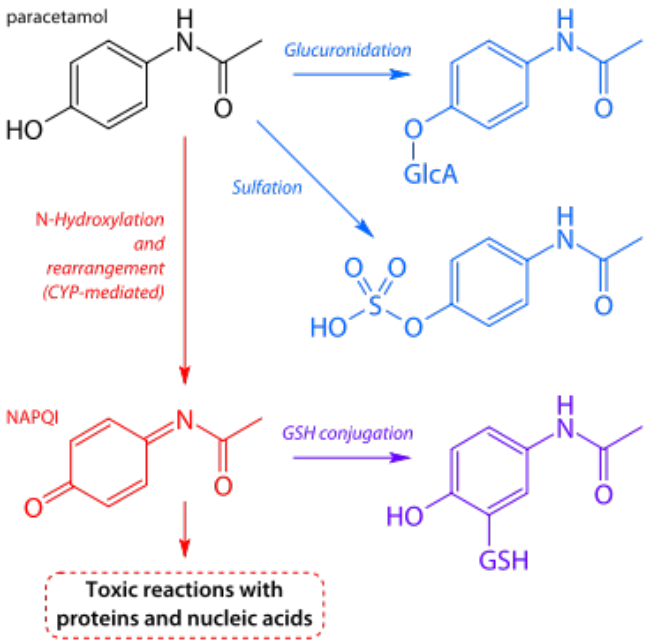


Quite a large number of metabolites have been reported for this compound. MetaPrint2D-React suggested locations for both of the reported aromatic hydroxylation reactions, and the glucuronidation, and predicted reduction of the nitro group. The model failed to predict the dephenylation reaction or the sulfonation, but did predict a number of transformations that have not been reported: dehalogenations, amide hydrolysis and an additional site of hydroxylation.

Molecule Count	% Top 1	% Top 3	Mean AUC	Median AUC
922	58.9%	78.7%	0.812	0.918

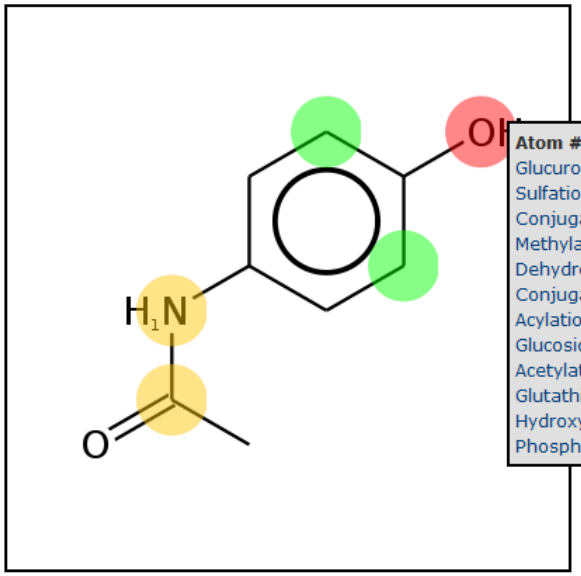
Table 19: Results of the evaluation of site of metabolism predictions made using MetaPrint2D-React.

Paracetamol toxicity (Tylenol) Overdose results in species NAPQI and liver damage



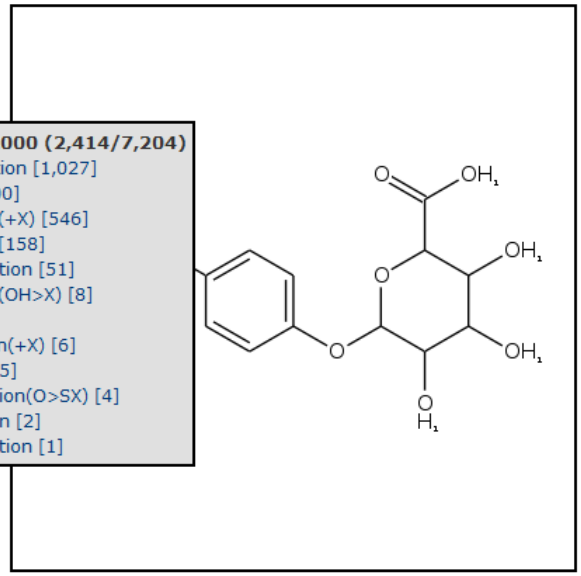
Metaprin2D results

Results



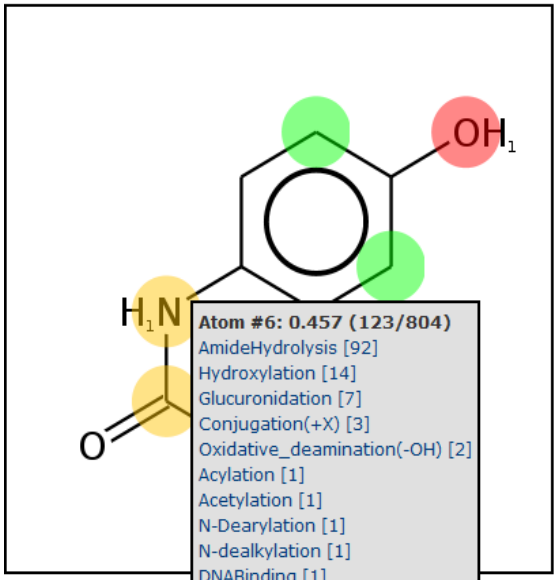
Reaction type filter: all

Metabolite



Reaction type: Glucuronidation

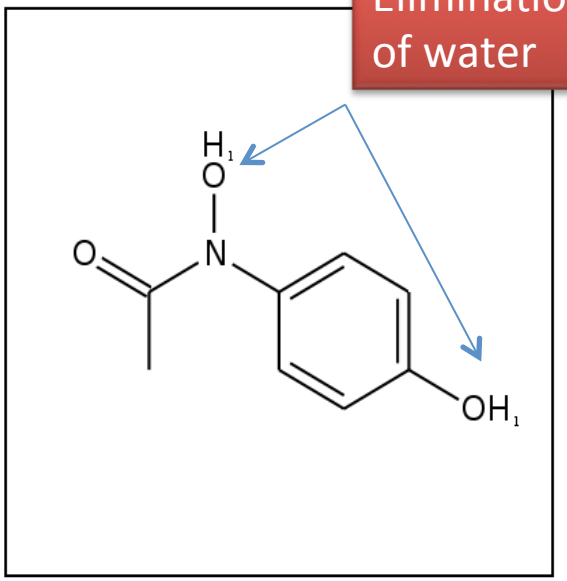
Results



Reaction type filter: all

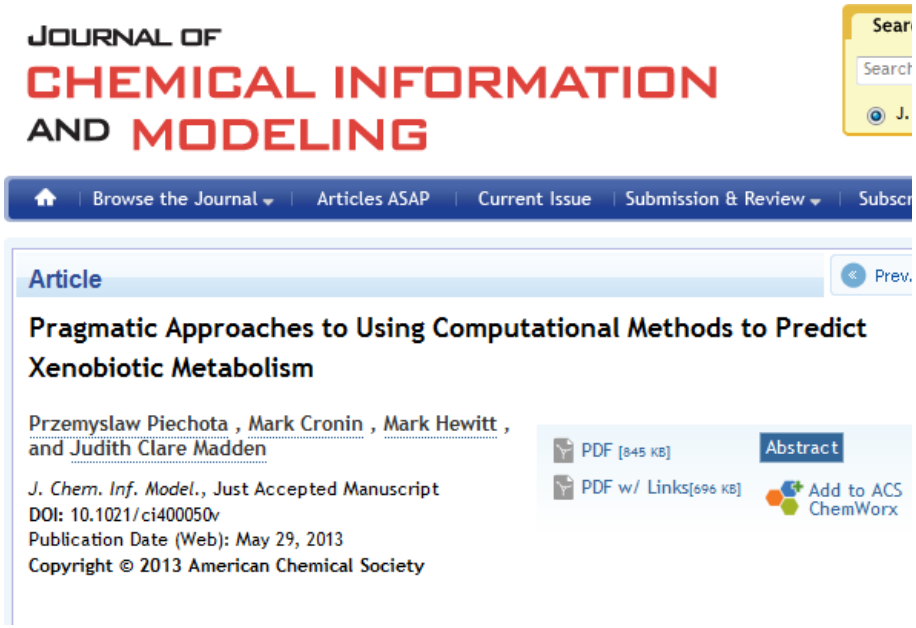
Metabolite

Elimination of water



Reaction type: Hydroxylation

Evaluation of metabolism prediction using Metaprint2D


JOURNAL OF CHEMICAL INFORMATION AND MODELING

Browse the Journal | Articles ASAP | Current Issue | Submission & Review

Article

Pragmatic Approaches to Using Computational Methods to Predict Xenobiotic Metabolism

Przemyslaw Piechota , Mark Cronin , Mark Hewitt , and Judith Clare Madden

J. Chem. Inf. Model., Just Accepted Manuscript
DOI: 10.1021/ci400050v
Publication Date (Web): May 29, 2013
Copyright © 2013 American Chemical Society

PDF [845 KB] Abstract
PDF w/ Links[696 KB] Add to ACS ChemWorx

Advantages: speed, comprehensive, coverage of metabolism, predicts products

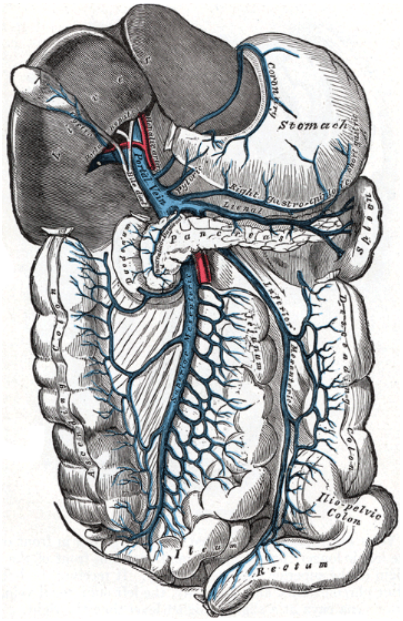
Disadvantages: relies totally on available data, restricted chemotypes, relies of reporting habits in journals. (more reported – more likely?)

Top_1 (Top_2) metabolites

Software	No. metabolites correctly predicted for DS1	No. metabolites correctly predicted for DS2
Meteor (setting EQU3)	57 (73%)	86 (85%)
Metaprint2D-React	62 (80%)	90 (89%)

DS1 – diverse set
DS2- Drugs

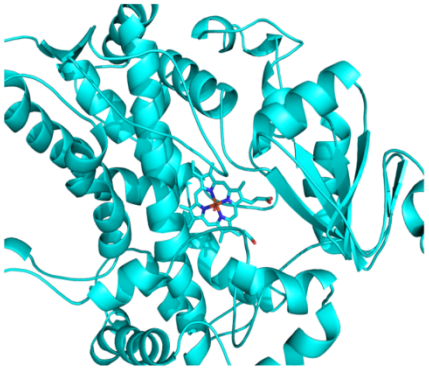
A SAR approach to Predicting sites of metabolism



•Many methods have been developed to predict **sites** of metabolism.
•Most focus on the main metabolism enzymes, the Cytochrome p450's
•These are ubiquitous, but in high concentration in the Liver, hence the focus on first pass metabolism of drugs as they are shuttled from the gut to the liver via the portal vein.

•They belong to a family of Iron containing enzymes, and are supreme oxidisers.
•However, they can also participate in other reactions e.g. in ring forming reactions to make steroids.

3A4



•The cytochrome P450s (CYPs) are a family of heme-containing enzymes involved in the phase-I metabolism of over 90% of drugs on the market . The CYP family of enzymes consists of 57 isoforms with the majority of biotransformations in mammals facilitated by the CYP 3A4 isoform, followed by 2D6 and 2C9.

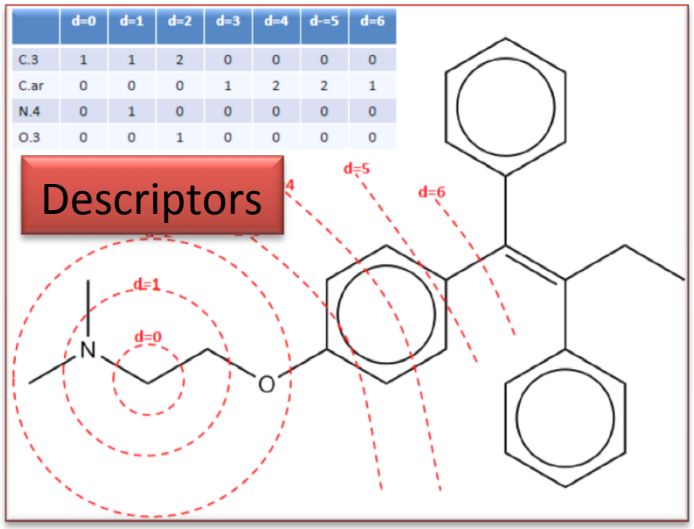
Cytochrome P450 site of metabolism prediction from 2D topological fingerprints using GPU accelerated probabilistic classifiers

Database of Cytochrome p450 substrates and products

2D topological fingerprints calculated to a bond depth of 4-6 contain sufficient information to allow the identification of SoMs using classifiers based on relatively small data sets
(1A2 -137 , 2C9-129, 2D6-157, 3A4-293, All-716)

Data Set	Top-2%						
	PRW (depth=6)	RASCAL	Xenosite [16]	RSPredictor [14,15]	SMARTCyp [8]	Reactivity & Docking [10]	Random
2C9	88	88	87	85	86	78	18
2D6	91	90	89	86	84	80	22
3A4	85	83	88	82	80	75	21

Table shows site of metabolism prediction results in terms of the top-2% compared to other methods.



Machine Learning. Parzen-Rosenblatt Window (PRW), Naive Bayesian(NB) and a novel approach called RASCAL (Random Attribute Sub-sampling Classification ALgorithm).

Prediction of metabolic sites in Cyps's.

Tyzack *et al. Journal of Cheminformatics 2014,* ***6:29***
http://www.jcheminf.com/content/6/1/29

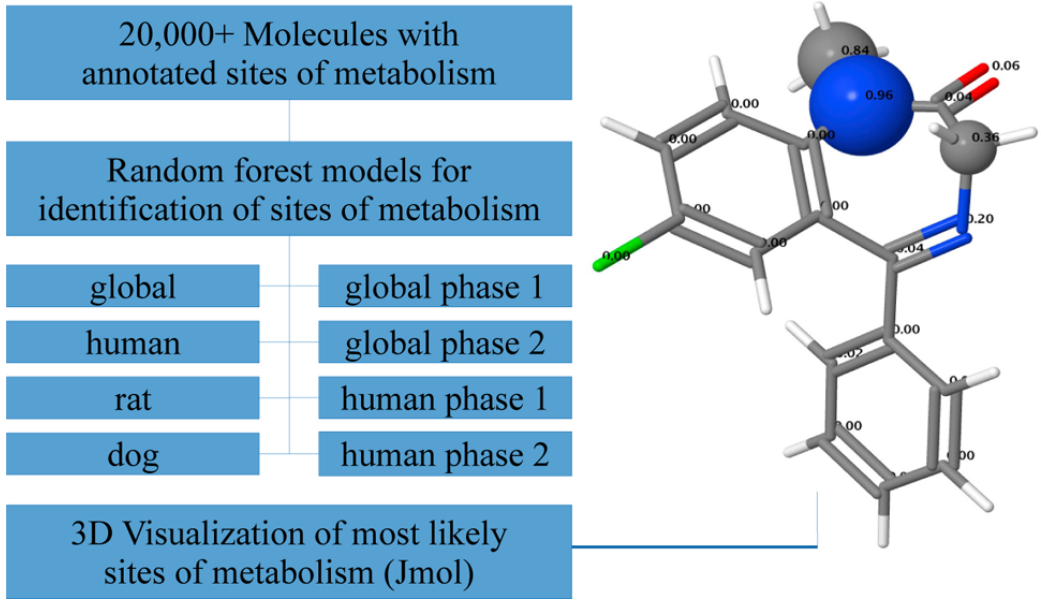
Journal of Cheminformatics

SAR and metabolism – using a large database, descriptors and machine Learning

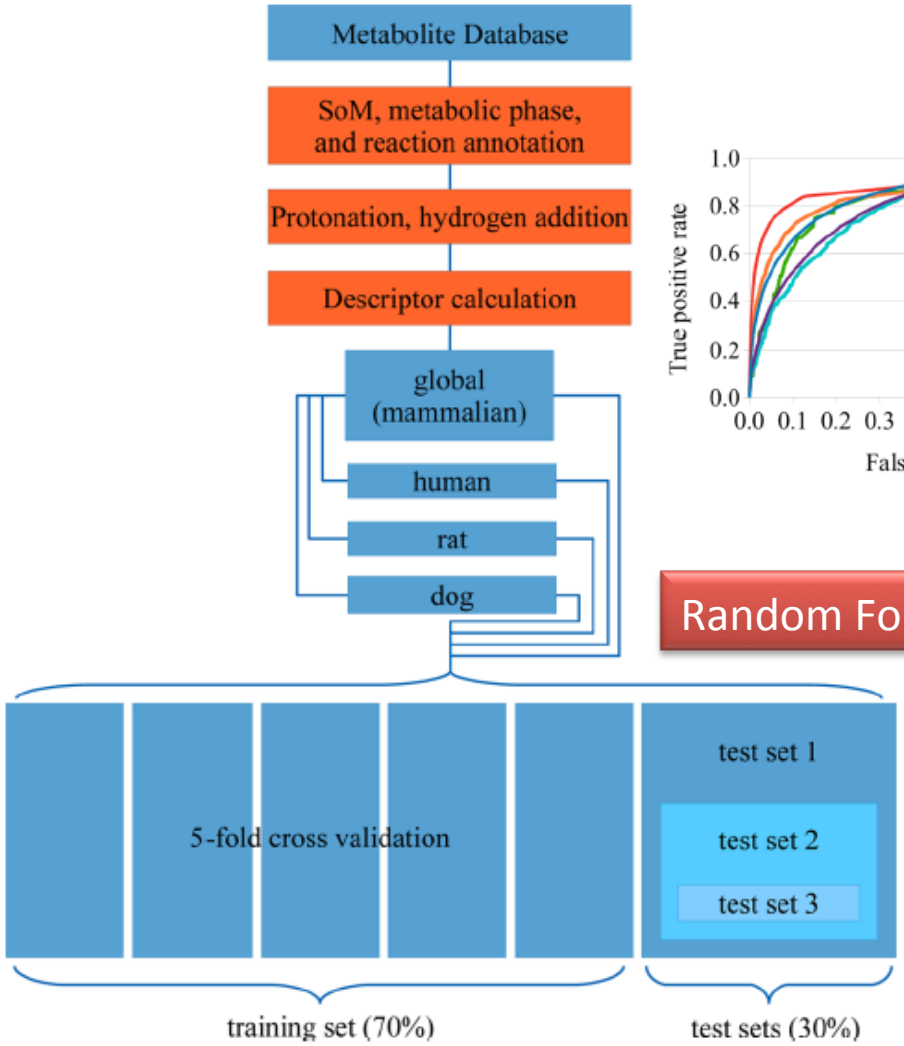
FAME uses a curated subset of the Metabolite database of 100,000 metabolic transformations

A global model of metabolism has been generated as well as specific models for human, rat, and dog metabolism. In addition, dedicated models are also available to predict SoMs of phase I and II metabolism.

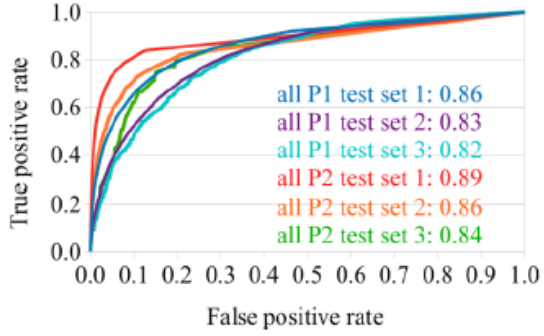
Only seven descriptors were found to be important (encoding the atom type and electronic configuration of each atom: total partial charges, sigma partial charges, pi electronegativity, sigma electronegativity, and polarizability) and one molecular descriptor encoding the topological size of a molecule.



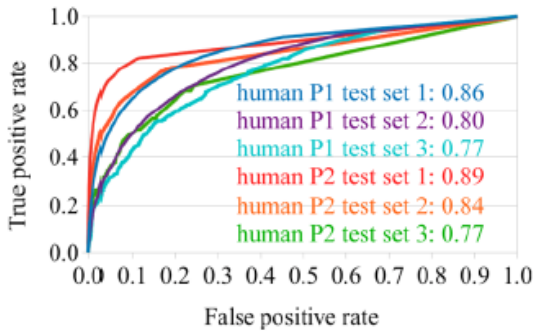
species	metabolic phase	no. mol[a] training set	no. mol[a] test set 1	no. mol[a] test set 2	no. mol[a] test set 3
all	phase 1 + 2	21098	9057	1889	181
human	phase 1 + 2	10347	4434	1065	149
rat	phase 1 + 2	13107	5622	1381	136
dog	phase 1 + 2	2929	1260	408	86



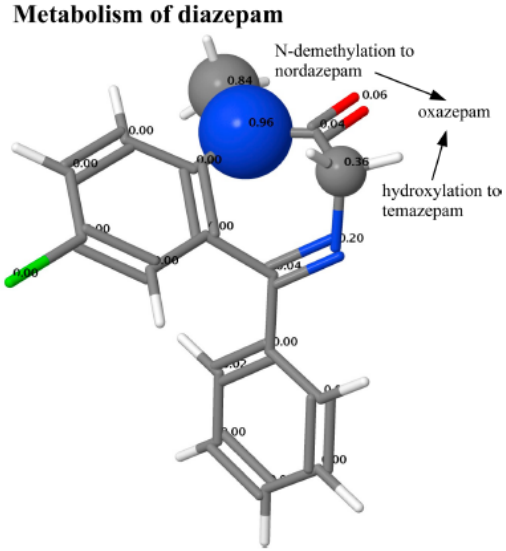
Random Forest models



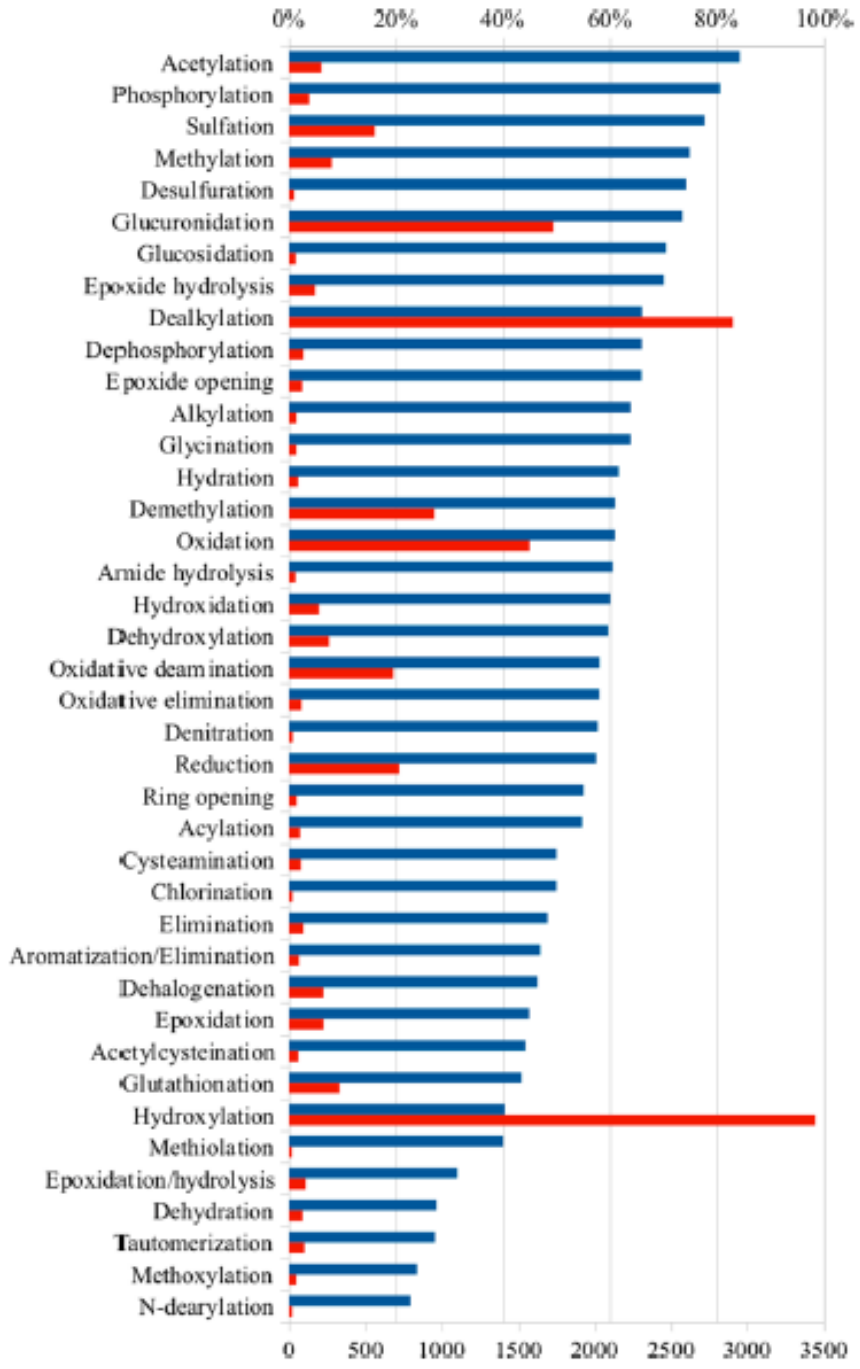
a



b



FAME is able to identify at least one known SoM among the top-1, top-2, and top-3 highest ranked atom positions in up to 71%, 81%, and 87% of all cases tested, respectively.



How common are reported metabolic transformations and how does the model perform?

Top-3 prediction rates for the global metabolism model on test set 1, itemized with respect to the reaction type. No correlation between propensity of a reaction type (the number of reactions: red bars) and prediction rates (blue bars) is observed.

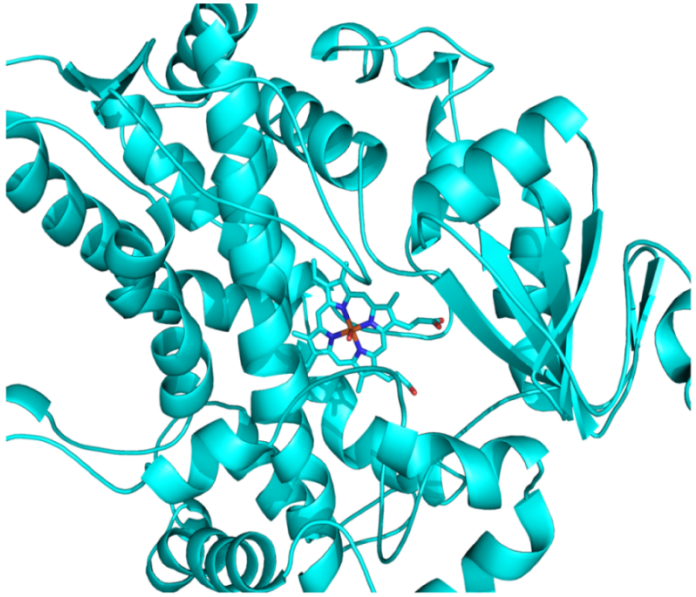
FAME is freely available from the authors to academia and nonprofit organizations.

FAst MEtabolizer (FAME): A Rapid and Accurate Predictor of Sites of Metabolism in Multiple Species by Endogenous Enzymes. Kirchmair et al., J. Chem. Inf. Model., 2013, 53 (11), pp 2896–2907.
DOI: 10.1021/ci400503s

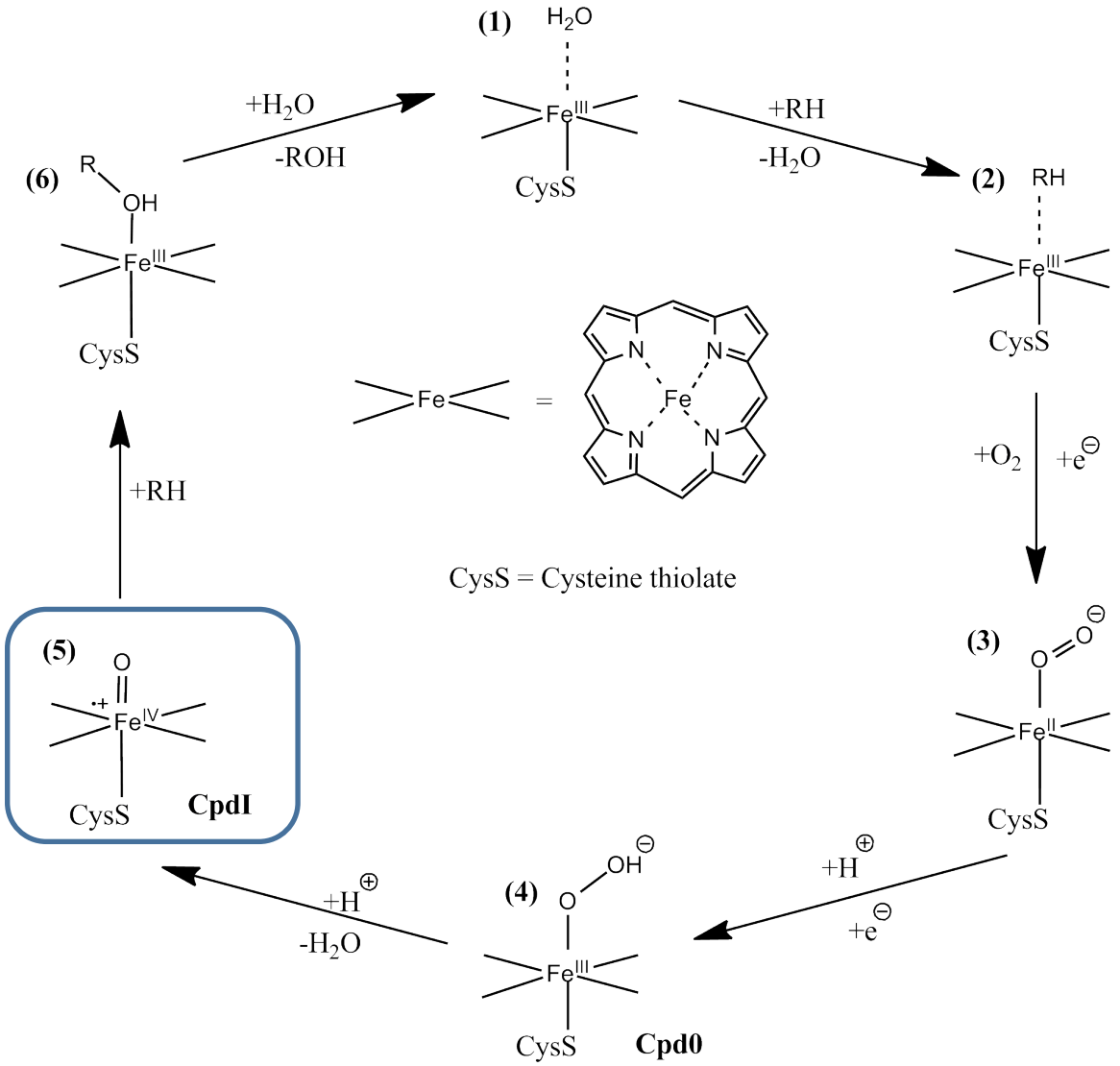
Prediction of sites of CYP450 Metabolism using Reactivity and Tethered Docking – a modelling approach

- CYP450 enzymes facilitate oxidation and bio-transformations
- Is a family of 57 similar haeme containing enzymes
- Most important are CYP3A4, CYP2D6, CYP2C8/9, CYP1A2
- Contain an iron-porphyrin reaction centre
- Metabolise 90% of drugs currently on the market

3A4



Cytochrome p450 Reactivity: Catalytic Cycle



Compound1 (Fe^{4+}) is the most reactive state. Reaction generally involves hydrogen abstraction or single electron transfer (SET) pathways

CYPs: Mechanism

Single electron transfer (SET) pathways

Aromatic C-oxidation

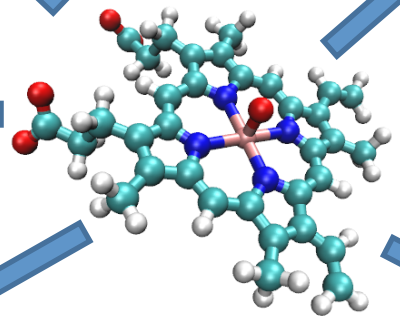
Ar-H ⟶ Ar-OH

Alkene epoxidation

C=C ⟶ C–O–C

S-oxidation
N-oxidation
P-oxidation

R_1–S(=O)–R_2 ⟶ R_1–S(=O)–R_2



Which site on the Molecule is preferentially Metabolised ? (SoM)

Hydrogen Abstraction

Aliphatic hydroxylation

R-H ⟶ R-OH

S-dealkylations
N-dealkylations
O-dealkylations

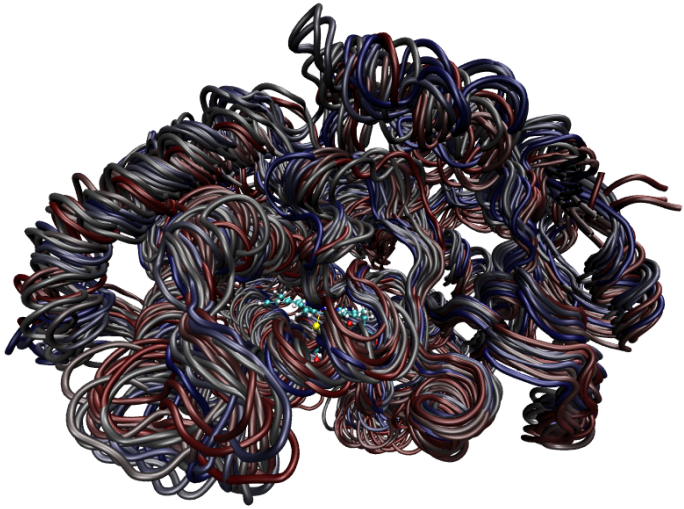
R_1-NH-R_2 ⟶ R_1-NH_2 + HO-R_2

SoM (Site of Metabolism) Prediction

- Need to consider reactivity and accessibility

  1. Accessibility – tethered docking with GOLD (a docking program)
  2. Reactivity – NWChem ground state molecular orbital analysis from DFT (Density Functional Theory)

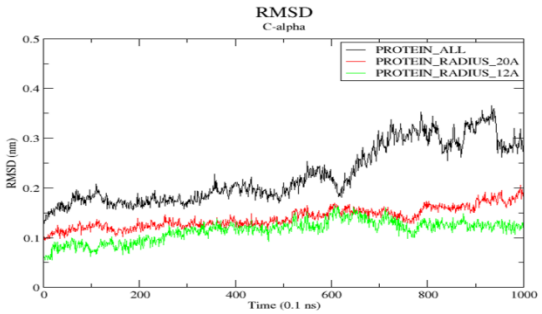
- Three isoforms considered:
  - 3A4 – 293 ligands
  - 2D6 – 157 ligands
  - 2C9 – 129 ligands

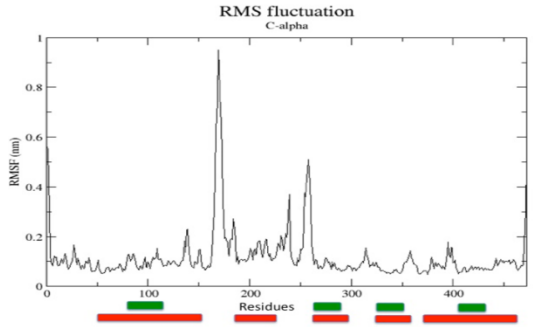


- Run Molecular Dynamics to optimise protein structures and select suitable frame(s) for docking. 300ns production MD simulation using AMBER (enzyme, solvent)

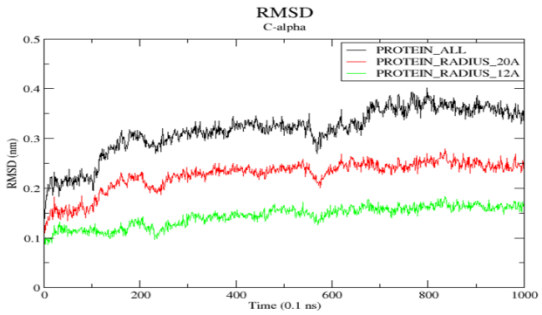
Molecular Dynamics, showing fluctuations in structure

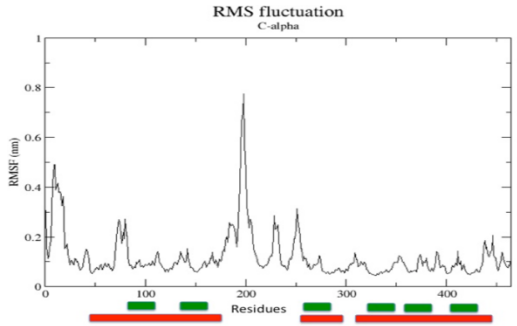
3A4



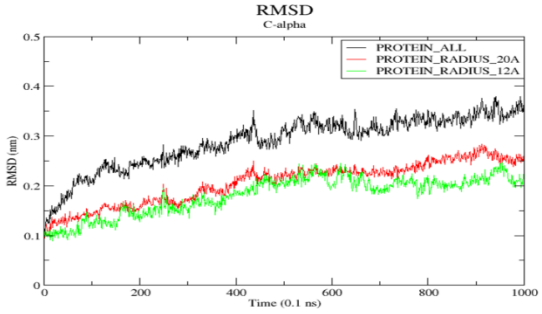


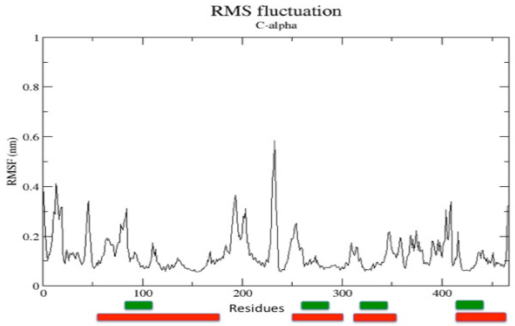
2D6

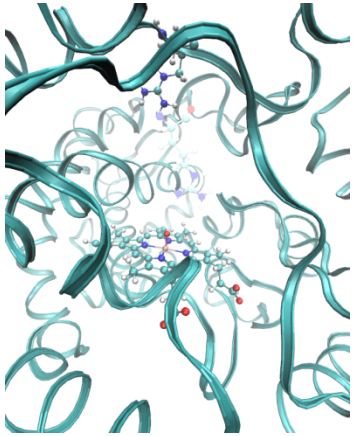




2C9

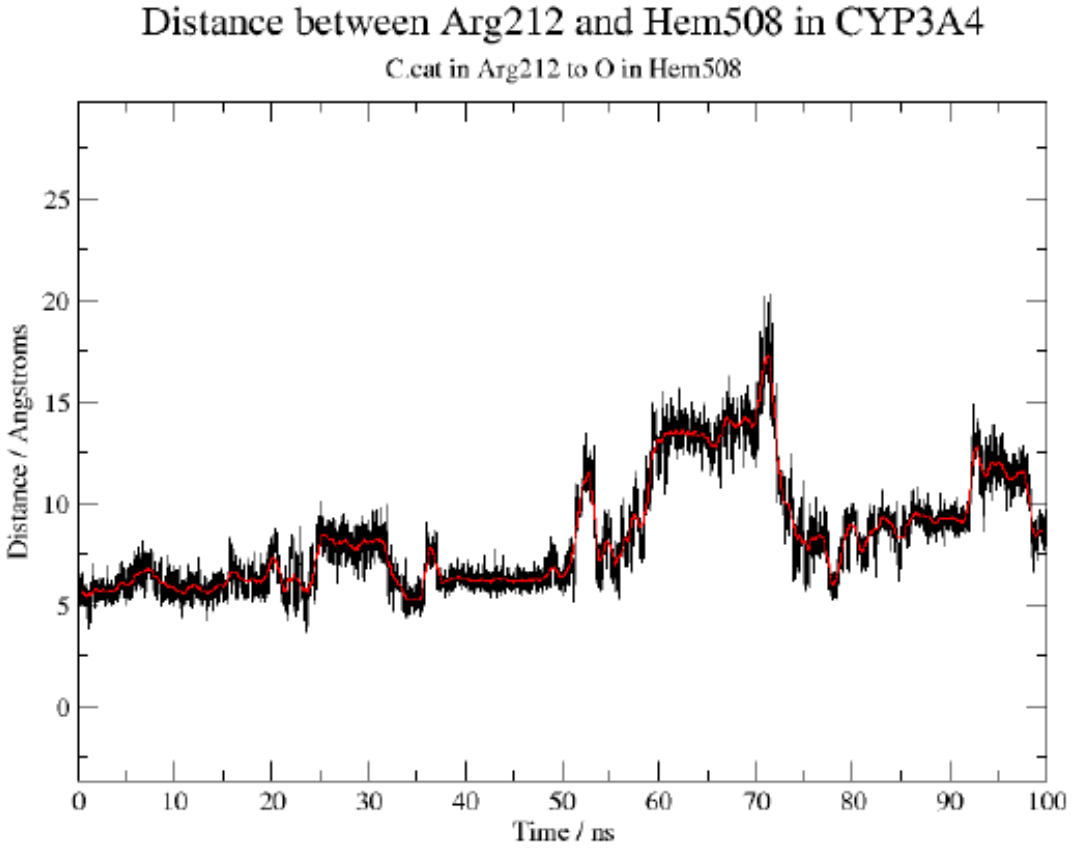


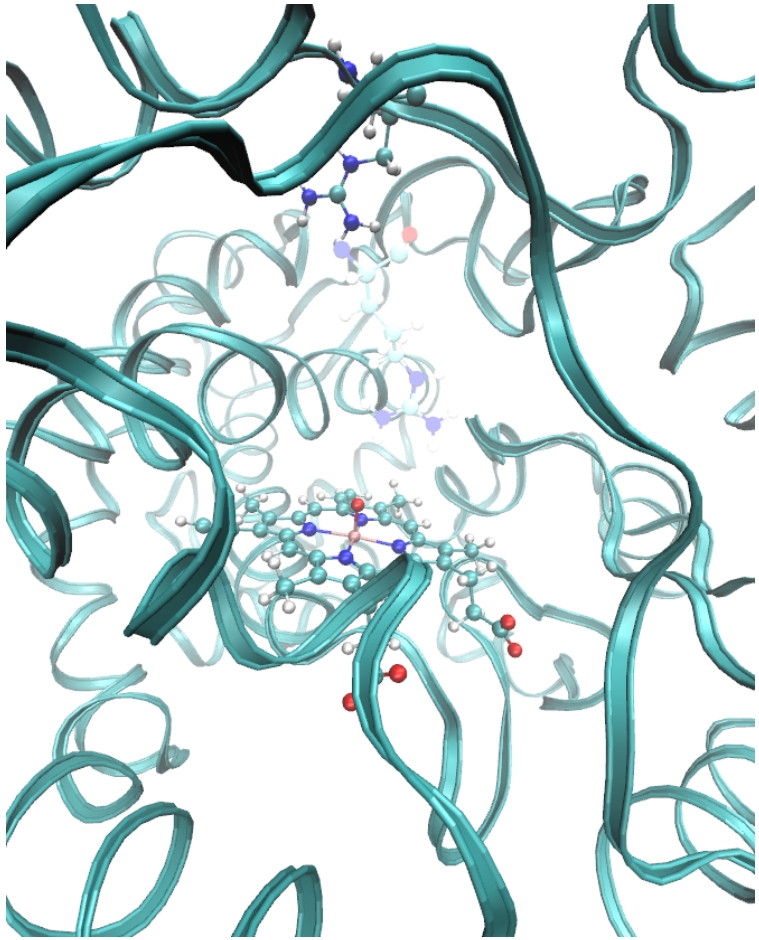




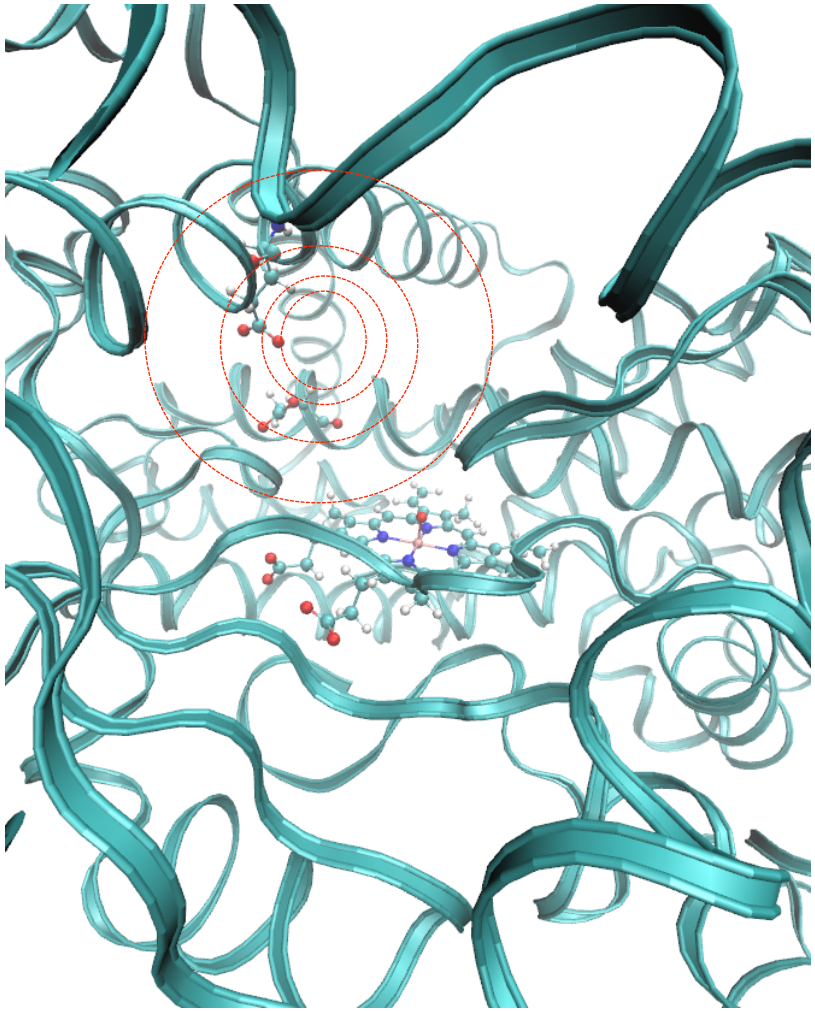
Some residues show large fluctuations in position –
We should take account of these

Flexible side-chains apparent from MD simulations. E.g. Arginine212 in Cyp-3A4 – this flexibility (and other flexible residues) could allow multiple docking modes





We have added a correction – to take into account the mobility of some charged residues – using position and mobility from the MD results. This correction is added to the docking score



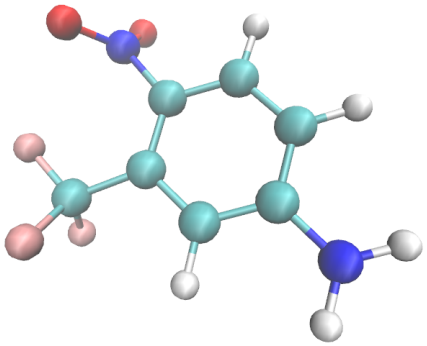
- Many ligands charged
- Charged residues around CYP cavity
- Use spherical constraints in GOLD to model electrostatics
- Calculate radius to model $1/r^2$ Coulombic attraction

Radius r Å	$\frac{1}{r^2} \times 10^3$	Incremental Score Boost	Cumulative Score Boost
4.000	62.50	0.5	5.0
4.216	56.25	0.5	4.5
4.472	50.00	0.5	4.0
4.781	43.75	0.5	3.5
5.164	37.50	0.5	3.0
5.657	31.25	0.5	2.5
6.325	25.00	0.5	2.0
7.303	18.75	0.5	1.5
8.944	12.50	0.5	1.0
12.649	6.25	0.5	0.5

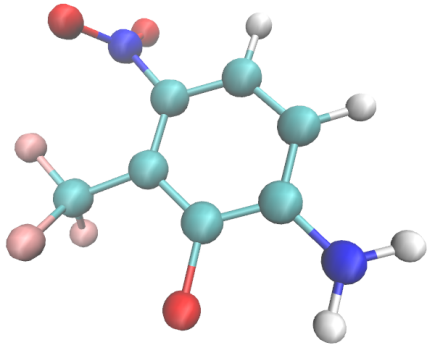
Calculating Accessibility of the ligand into the active site: Tethered Docking

- Tether each potential SoM to CpdI and use GOLD docking score to rank
- Replace each H in turn with an O and use as dummy link atom to reactive O in heme

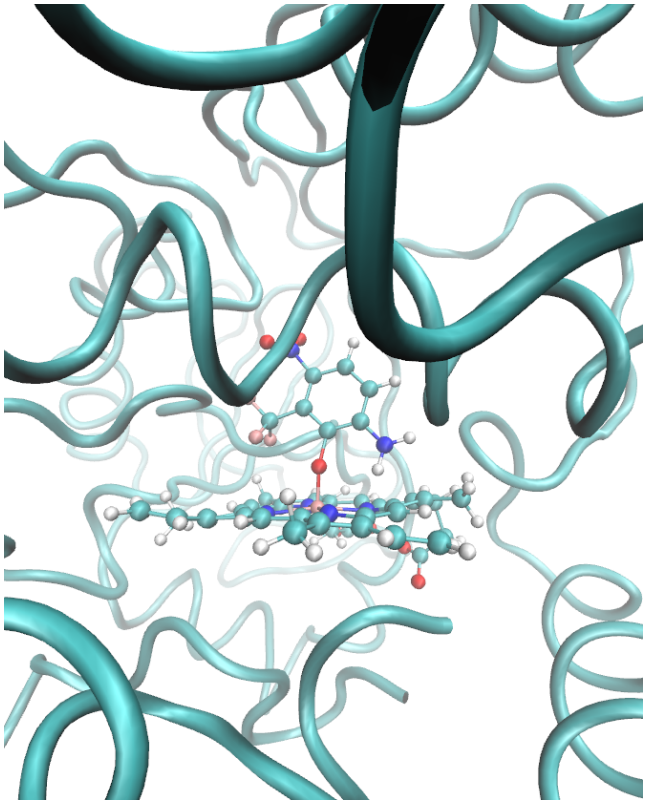
I.



II.



III.



Accessibility: docking results. We rank predictions as finding the 'top-2' or 'top-3' metabolism sites from experiment

Scoring Function	Parameter Set	Tether Length Å	3A4 Top 3 %	3A4 Top 2 %	2D6 Top 3 %	2D6 Top 2 %	2C9 Top 3 %	2C9 Top 2 %
Chemscore	PDB	2.0	34	24	47	36	61	48
Chemscore	CSD	2.0	33	23	49	40	61	45
Chemscore	default	2.0	37	27	54	41	62	47
Goldscore	PDB	2.0	49	34	66	53	71	61
Goldscore	CSD	2.0	50	37	64	50	73	60
Goldscore	default	2.0	50	34	65	54	70	59
ASP	default	2.0	59	44	61	50	66	61
PLP	default	no tether	48	35	46	37	59	47
PLP	default	2.0	68	56	74	63	72	64
PLP	**default**	**1.5**	**69**	**55**	**76**	**67**	**73**	**65**
PLP + constraints	default	no tether	39	28	48	35	62	43
PLP + constraints	**default**	**1.5**	**69**	**55**	**77**	**73**	**74**	**67**
PLP + constraints	default	2.0	63	50	75	67	74	68
PLP + constraints	default	2.5	55	39	70	55	73	62
PLP + constraints	default	3.0	44	33	61	48	74	63

Scoring function PLP and tethered docking at 1.5A Is best

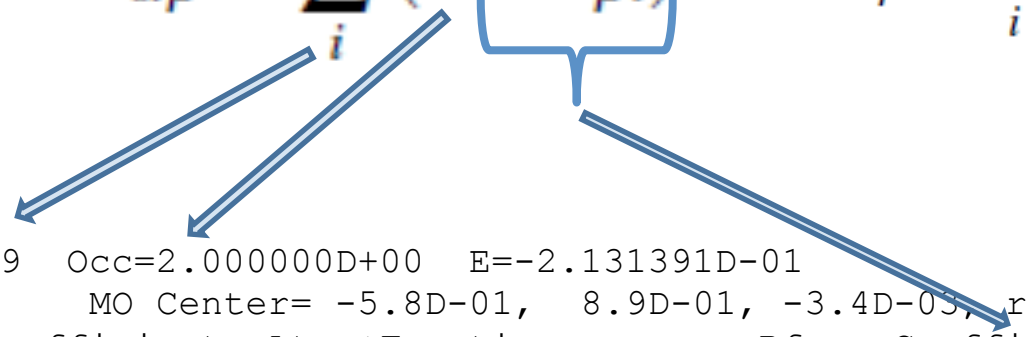
Adding reactivity: Using Molecular Orbital Analysis

- Hydrogen bond order is a good indicator of hydrogen abstraction
- Relevant to aliphatic hydroxylation, and subsequent dealkylation
- Also need a reactivity measure for electron abstraction which allows direct comparison to the hydrogen abstraction measure
- Reactivity measure derived from ground state molecular orbital analysis
- Geometry optimisation with NWChem using DFT, B3LYP, 6-31g**

Reactivity: Bond Order derived from the Density Matrix to give an energy weighted reactivity index

- Density Matrix defined as:

$$D_{\alpha\beta} = \sum_{i}^{MO} \left(n_i c_{\alpha i} c_{\beta i} \right) \qquad D^{w}_{\alpha\beta} = \sum_{i}^{MO} \left(n_i c_{\alpha i} c_{\beta i} (\varepsilon_i - \varepsilon_T) \right)$$



```
Vector    79  Occ=2.000000D+00  E=-2.131391D-01
                MO Center= -5.8D-01,  8.9D-01, -3.4D-03, r^2= 9.0D+00
   Bfn.  Coefficient  Atom+Function          Bfn.  Coefficient  Atom+Function
  ----- ------------  ---------------       ----- ------------  -------------
    108     -0.236743   8 C  px              202     -0.223507  14 C  px
    142      0.222353  10 C  px              206     -0.187117  14 C  px
    112     -0.184346   8 C  px              157      0.157935  11 C  px
    146      0.157431  10 C  px              109      0.145952   8 C  py
    143     -0.140489  10 C  py              172     -0.140463  12 C  px
```

Reactivity: Energy Weighted Density Matrix

- Starting with the density matrix and overlap matrix, the approach adopted here is to weight contributions to the density matrix by the molecular orbital energy to create an energy weighted density matrix and then calculating a weighted bond order.
- It has been shown that the total bond order of a hydrogen atom is a good indicator of SoM since it is indicative of bond strength and thus the ease of hydrogen abstraction. However, in order to extend this approach to be able to predict single electron transfers (SET) it is insufficient to only use information from the density and overlap matrices.
- Because of differences in valency of atoms and multiple bond types, we correct for this. This is achieved by calculating a reactivity score per unit bond order. The reactivity scores are added until the equivalent of one electron is reached.
- In this way reactivity scores for hydrogen and electron abstractions can be obtained from the same data source enabling the two competing pathways for CYP metabolism to be compared.

Reactivity: Results

Table 5: Reactivity results

Method	Energy Translation (ε_T)	3A4		2D6		2C9	
		Top 3 %	Top 2 %	Top 3 %	Top 2 %	Top 3 %	Top 2 %
Density matrix		70	60	55	45	73	66
Energy-weighted density matrix	50	77	65	71	58	77	57
Energy-weighted density matrix	100	78	67	65	53	78	60
Energy-weighted density matrix	150	79	67	65	52	78	63
Energy-weighted density matrix	200	78	68	64	53	79	64

Manuscript in preparation

Overall: Results

Table 6: Overall results versus comparatives from IMPACTS.

Method	3A4 [a]		2D6 [b]		2C9 [c]	
	Top 3 %	Top 2 %	Top 3 %	Top 2 %	Top 3 %	Top 2 %
This work	84	71	80	76	80	
IMPACTS	-	72-75	-	76	-	

$$Score_A = \frac{Reactivity}{Docking^P}$$

Advantages: Is clearly not limited by availability of metabolism data as this model is based on a simulation

Prediction of Cytochrome P450 Xenobiotic Metabolism: Tethered Docking and Reactivity Derived from Ligand Molecular Orbital Analysis
Jonathan D. Tyzack, Mark J. Williamson, Rubben Torella, and Robert C. Glen
J. Chem. Inf. Model., 2013, 53 (6), pp 1294–1305.
DOI: 10.1021/ci400058s

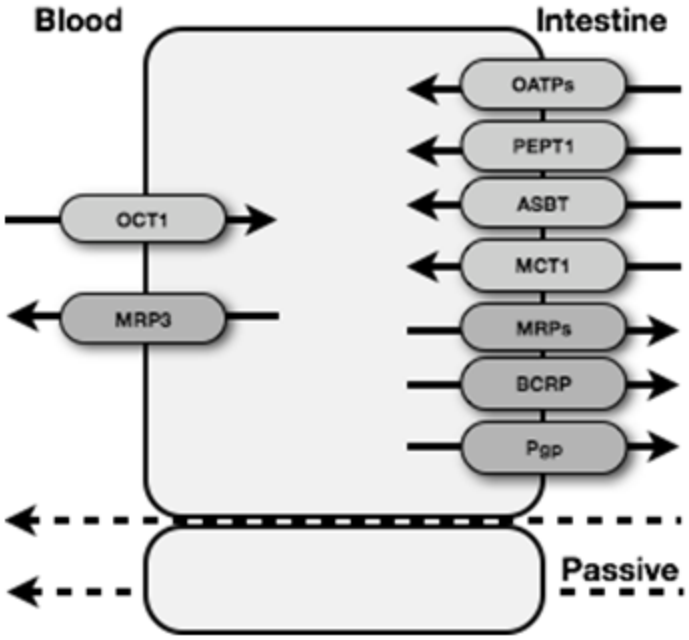
Transporters: moving molecules to the right place a the right time

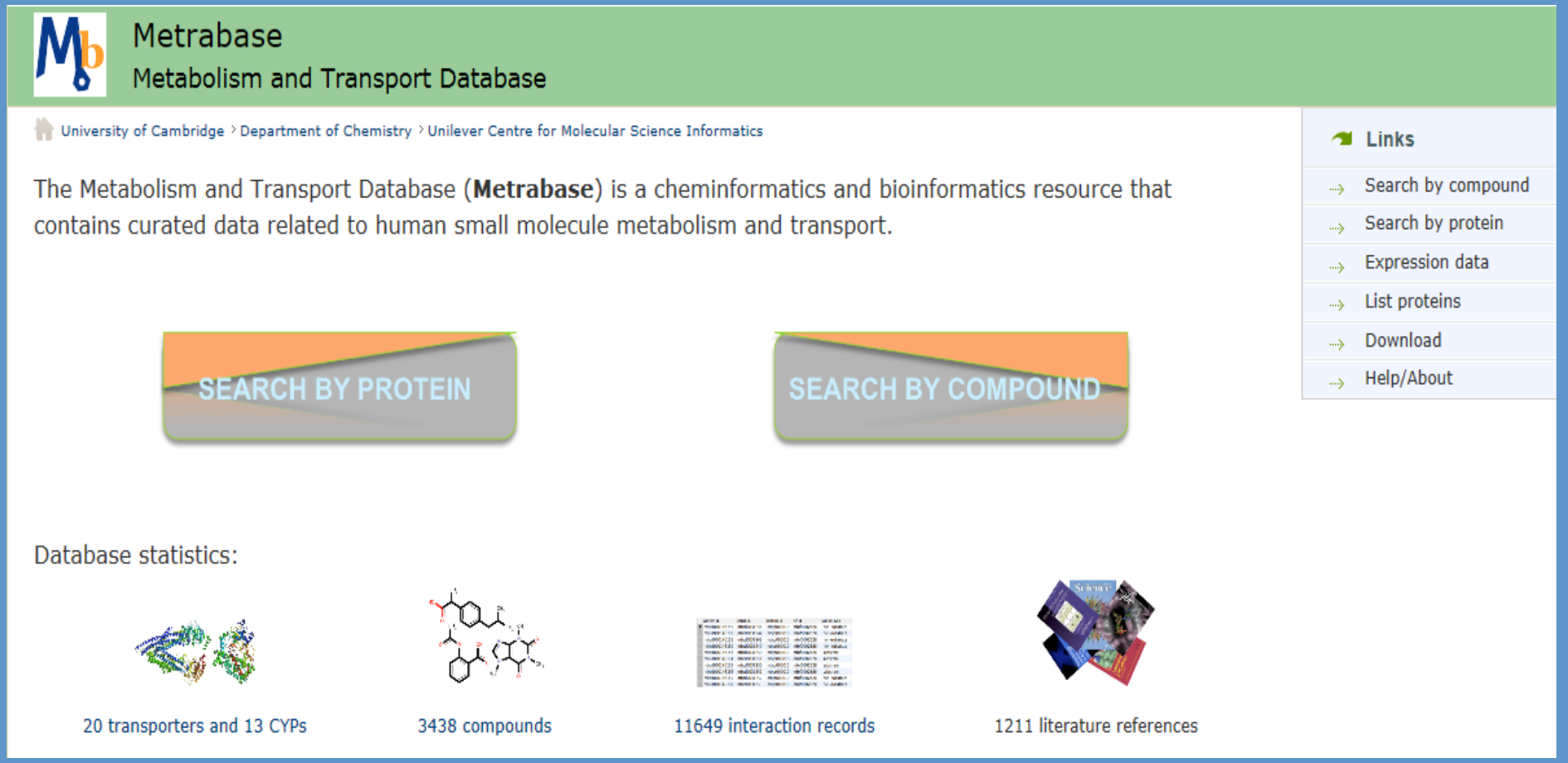
Transporters are responsible for the movement of most small, water-soluble, organic molecules and some inorganic ions across cell membranes. Each transporter is thought to be highly selective, often transferring just one type of molecule – but recent evidence points to more promiscuous behaviour.

Many drugs act at transporters e.g. SERT is responsible for the reuptake of extracellular serotonin (5-HT). E.g Prozac (41 cpds. at least act at SERT)

Transporters can
- be passive, moving a molecule along an electrochemical gradient e.g. glucose transporter
- Pumps actively transport a solute against an electrochemical gradient – may be ATP-driven, Light driven or coupled (in and out) e.g. Na^+ out and K^+ in
- Coupled pumps – symporters move substrates in one direction, antiporters move them in opposite directions. E.g. NCKX (sodium, calcium, potassium antiporter)

- We have constructed a database of Transporters





Transporter	Substrates	Non-Substrates	Inhibition	Induction
ABCB1	571	479	385	55
ABCG2	309	197	632	17
SLC15A1	247	95	294	6
ABCC2	160	139	284	116
SLC22A1	167	94	331	3
ABCC1	98	91	8	0
SLCO1B1	99	39	375	6
SLCO2B1	48	73	297	23
SLCO1B3	59	35	297	9
ABCC3	68	22	49	21
SLCO1A2	56	22	51	1
SLC10A2	54	19	11	0
ABCC4	47	19	0	0

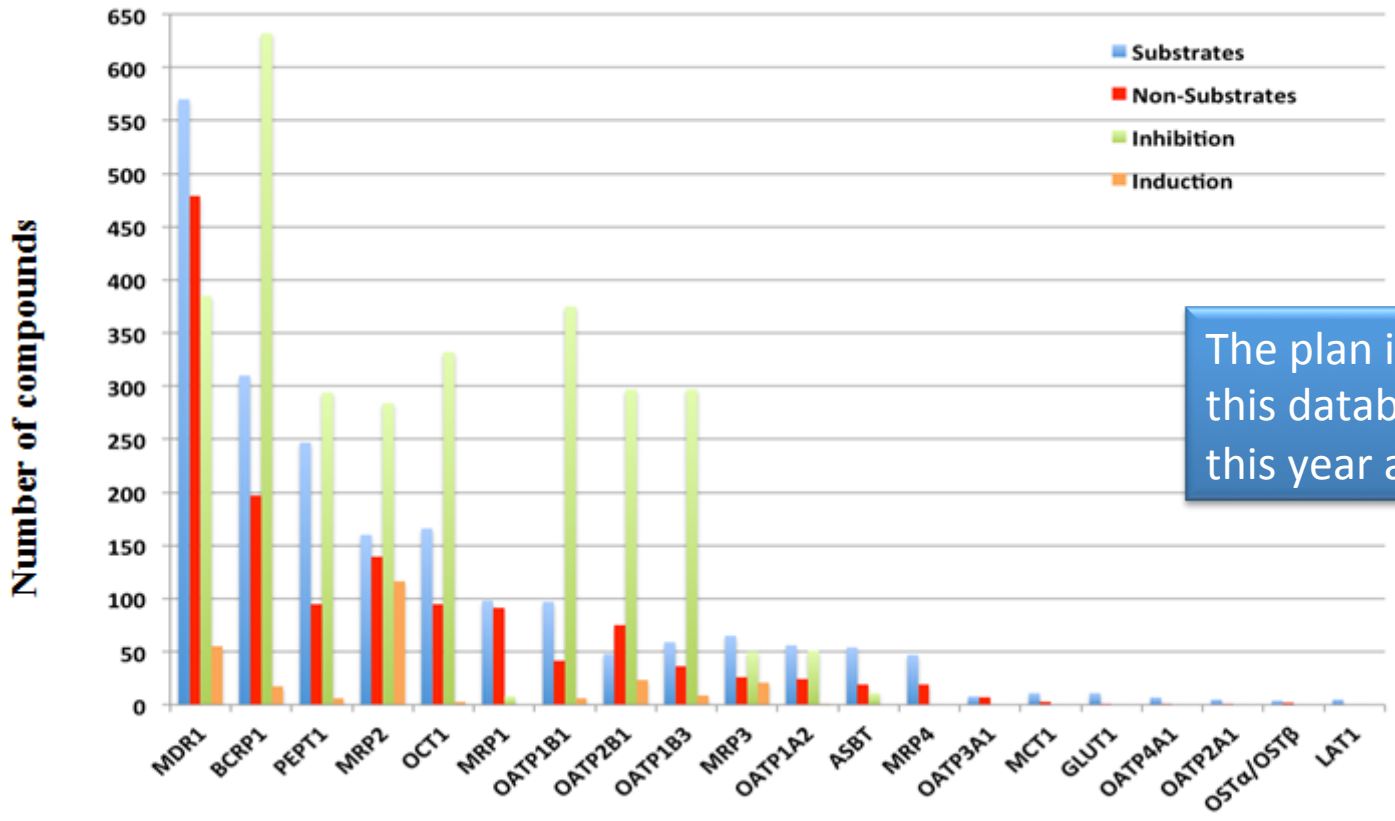
Metrabase is an integrated cheminformatics and bioinformatics resource containing curated data related to human transport and metabolism of chemical compounds. Its primary content includes over 3000 small molecule substrates and modulators of transport proteins and, currently to a smaller extent, cytochrome P450 enzymes.

MetraBase

Metrabase v1.0
20 transporters and 13 CYPs: 3438 compounds, 11649 interaction records, 1211 literature references

13 CYPs: 212 compounds, 506 interaction records, 36 literature references

A summary of the transporter-related Metrabase content



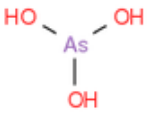
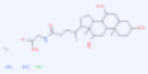
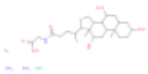
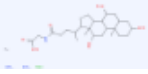
The plan is to make this database available this year as open source

Example search for substrates of the OATP1B1 transporter

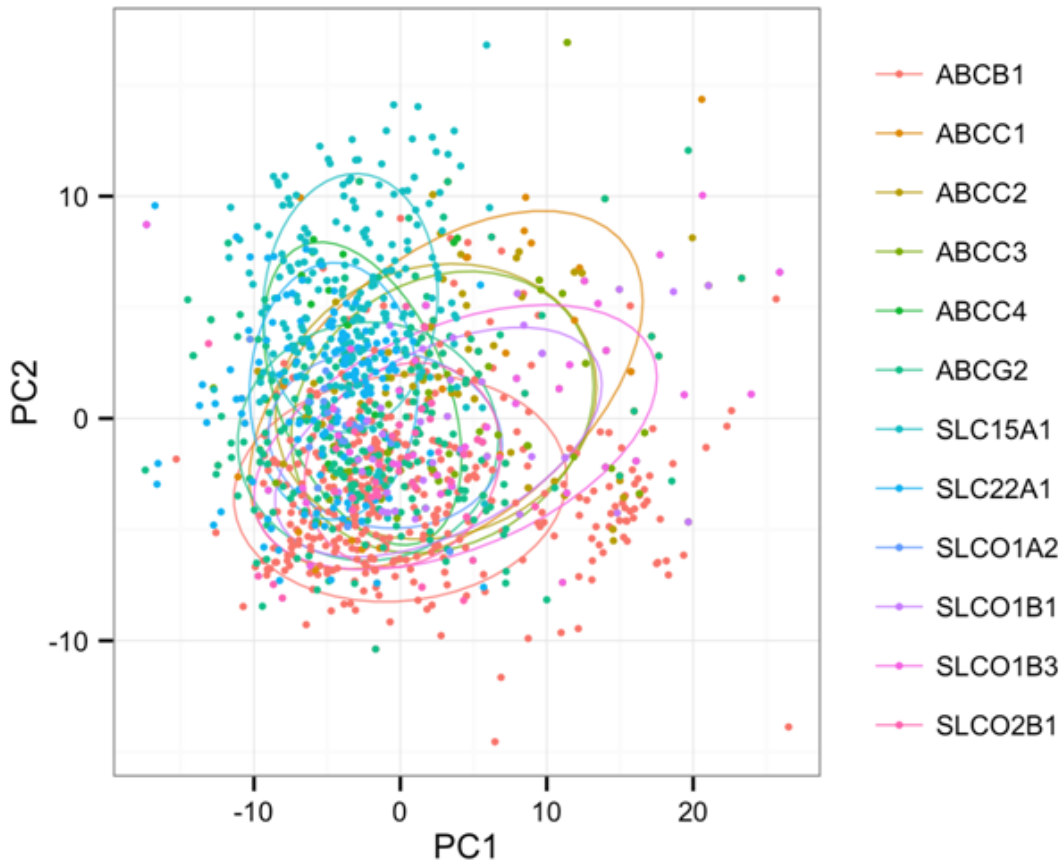
Results

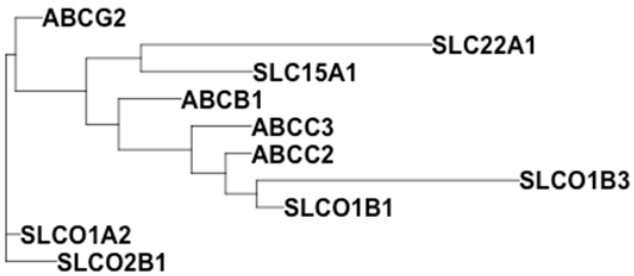
339 hits (97 compounds)

Download TSV file Download SD file

	Name	Action	Reference	Data source
	[D-Ala2, D-Leu5]-enkephalin	substrate Cell system: Xenopus laevis oocytes	PMID: 21245207 DOI: 10.1124/pr.110.002857 Abstract Cmpd: DADLE	literature
	Arsenite	substrate Cell system: HEK293	PMID: 21245207 DOI: 10.1124/pr.110.002857 Abstract Cmpd: arsenite	literature
	Bamet-R2	substrate Km = 10 µM Cell system: Xenopus laevis oocytes	PMID: 11901224 DOI: 10.1124/mol.61.4.853 Abstract Cmpd: Bamet-R2	TP-search
	Bamet-R2	substrate Km = 10 µM Cell system: Xenopus laevis oocytes	PMID: 17574004 DOI: 10.1016/j.beem.2007.03.004 Abstract Cmpd: bamet-R2	literature
	Bamet-R2	substrate Vmax/Km = 450 nL oocyte-1 (10 min)-1 Cell system: Xenopus laevis	PMID: 11901224 DOI: 10.1124/mol.61.4.853	literature

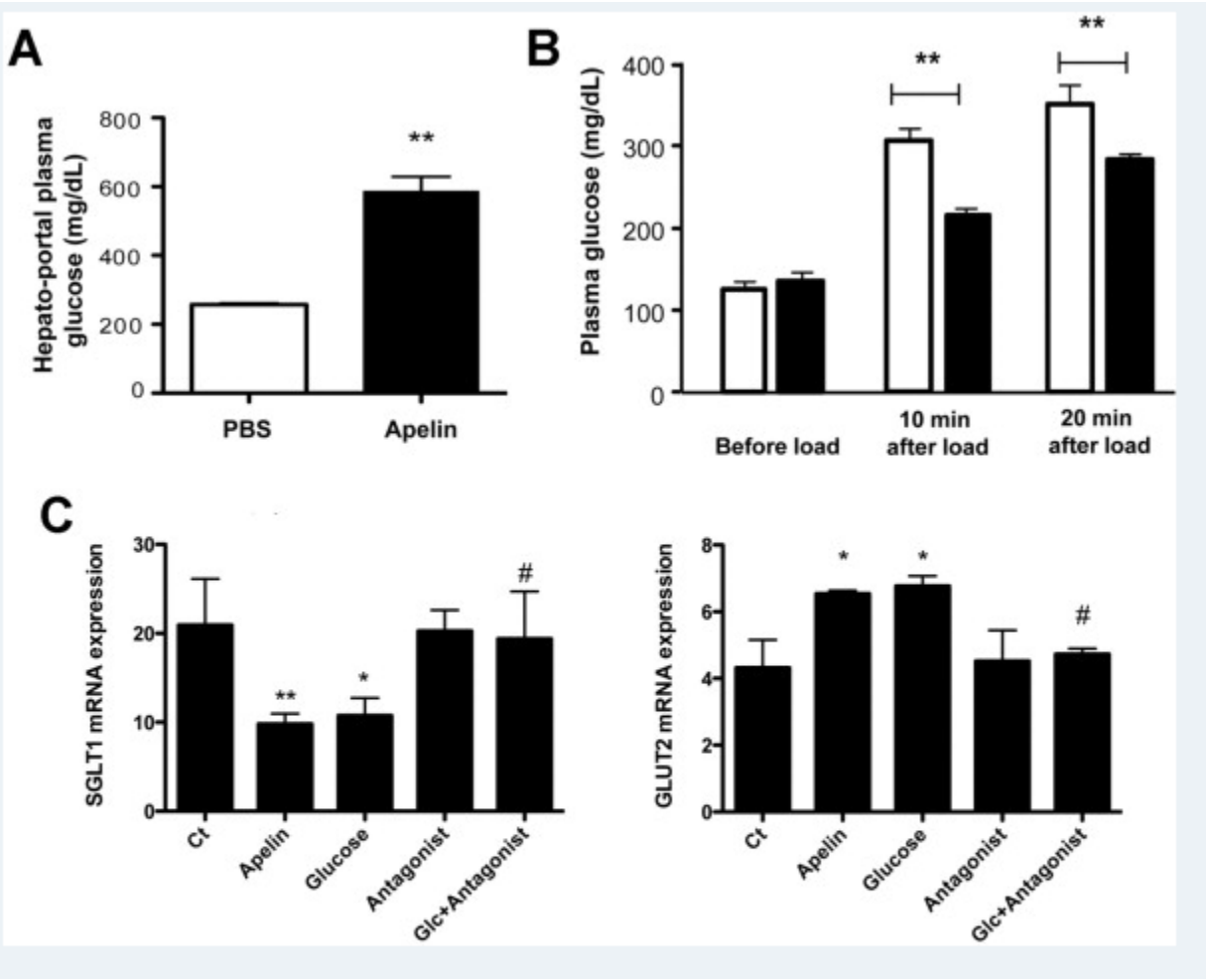
“Transporter Space”

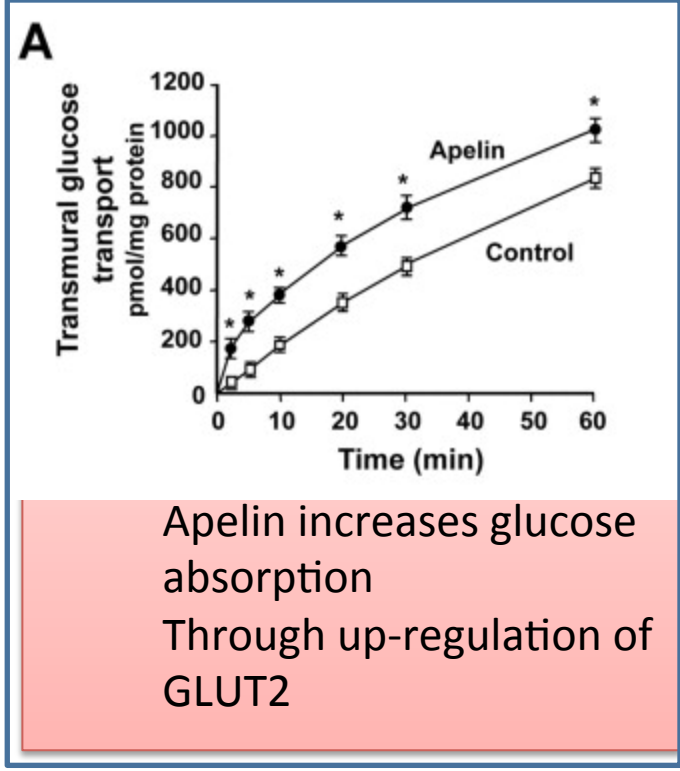




Cluster analysis

PCA based on 2D descriptors of the transporter substrates dataset, circles show the 80% approximate coverage in each transporter. Although SLC15A1 (peptide transporter) and Pgp substrates differ substantially by property (red vs. light blue) the majority of efflux and uptake transporters overlap significantly. Not as specific as thought!





Apelin increases glucose absorption
Through up-regulation of GLUT2

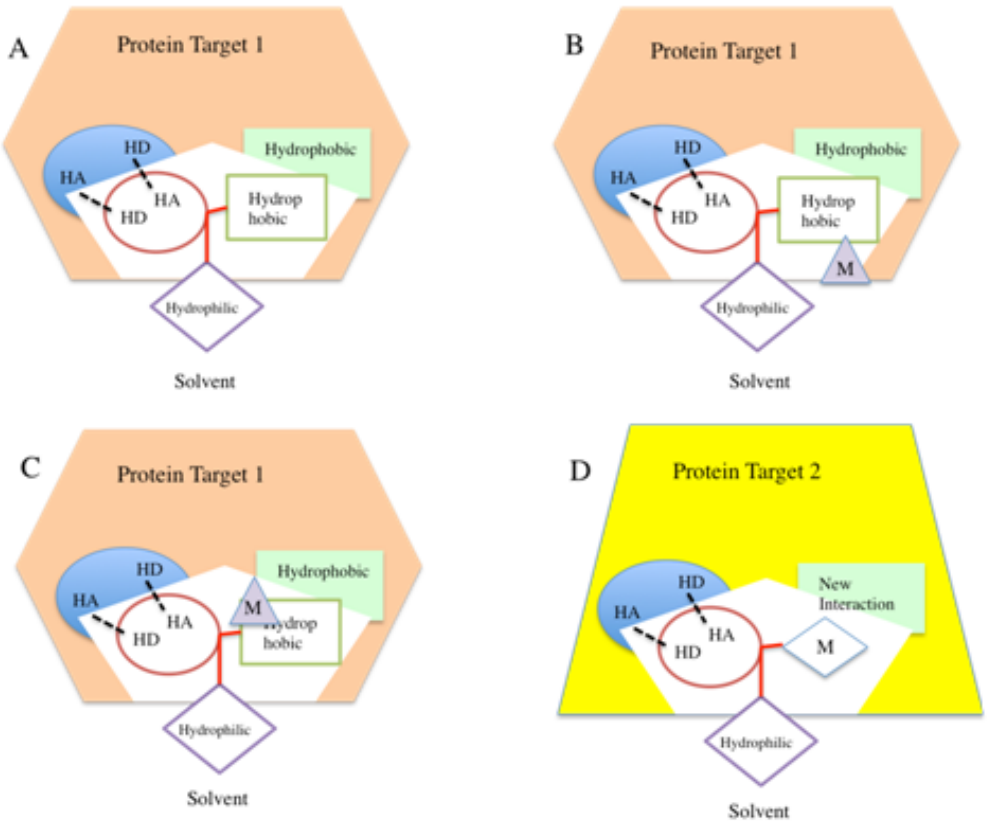
Orally administered apelin Increases antestinal transepithelial glucose transport from lumen to bloodstream in mice. The antagonist MM54 blocks this.

Gastroenterology. 2013 Apr;144(4):771-80. doi: 10.1053/j.gastro.2013.01.004. Epub 2013 Jan 10.

The intestinal glucose-apelin cycle controls carbohydrate absorption in mice.

Dray C, Sakar Y, Vinel C, Daviaud D, Masri B, Garrigues L, Wanecq E, Galvani S, Negre-Salvayre A, Barak LS, Monsarrat B, Burlet-Schiltz O, Valet P, Castan-Laurell I, Ducroc R.

Modification of activity by metabolism



Drug metabolism redox reactions such as heteroatom dealkylations, hydroxylations, heteroatom oxygenations, reductions, and dehydrogenations can yield active metabolites, and in rare cases even conjugation reactions can yield an active metabolite.

Obach, R. S., Pharmacologically active drug metabolites: impact on drug discovery and pharmacotherapy. *Pharmacological reviews* **2013,** *65* (2), 578-640.

Metabolism and pro-drugs –more common than you might think

A
- Acemetacin
- O-Acetylpsilocin
- Aconiazide
- Adrafinil
- Alatrofloxacin
- Aldophosphamide
- Amfecloral
- Amifostine
- Amlodipine/benazepril
- Ampiroxicam
- Arbaclofen placarbil
- Avizafone
- Azathioprine

B
- Bacampicillin
- Bambuterol
- Benazepril
- Bezitramide
- BMS-663068
- Bopindolol
- Brincidofovir
- Brivanib alaninate
- 1,4-Butanediol

C
- Capecitabine
- Carbamazepine
- Carfecillin
- Carindacillin
- Carisoprodol
- Cefuroxime axetil
- Chloral hydrate
- Clobenzorex
- Clofibrate
- Clofibride
- Cloforex
- Clopidogrel
- Cloxazolam
- Codeine
- Codrug
- Combretastatin A-4 phosph
- CRL-40,941
- Cyclophosphamide
- Cyprodenate

D
- Deflazacort
- Delapril
- Dextromethorphan
- DHA-clozapine
- Dipivefrine
- Dirithromycin
- Dolasetron
- L-DOPA
- Droxidopa

E
- Enalapril
- Etilevodopa
- Etofibrate

F
- Famciclovir
- Fenethylline
- Fenofibrate
- Fesoterodine
- Fimasartan
- Fosamprenavir
- Fosaprepitant
- Fosfluconazole
- Fosinopril
- Fosphenytoin
- Fospropofol
- Fostamatinib
- Fursultiamine

G
- Gabapentin enacarbil
- Glycerol phenylbutyrate

H
- Heroin
- Hetacillin

I
- Ibotenic acid
- Indometacin farnesil
- Irinotecan
- Isoniazid

L
- Levomethorphan
- Lisdexamfetamine
- Loxoprofen

M
- Melevodopa
- Mestranol
- Methyl aminolevulinate
- Midodrine
- Moexipril
- 6-Monoacetylmorphine

N
- Nabumetone
- Nitazoxanide

P
- PA 824
- Parecoxib
- Pirisudanol
- Pivampicillin
- Pivmecillinam
- Potassium canrenoate
- Prednisone
- Proglumetacin
- Proguanil
- Prontosil
- Pyrazinamide

Q
- Quinapril

R
- Rabeprazole
- Ramipril
- Rilmazafone
- Romidepsin
- Ronifibrate

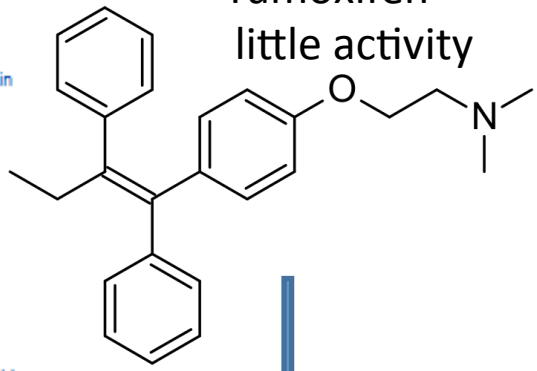
S
- Sergliflozin etabonate
- Sibrafiban
- Sodium phenylbutyrate
- Sofosbuvir
- Spirapril
- Spiruchostatin
- Sulindac

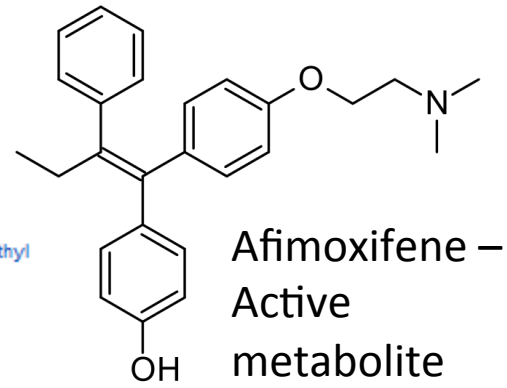
T
- Tamoxifen
- Taribavirin
- Tebipenem
- Tegafur
- Temocapril
- Temozolomide
- Tenatoprazole
- Tenofovir
- Terfenadine
- Tramadol
- Trandolapril
- Triclofos
- Tybamate

V
- Valaciclovir
- Valganciclovir
- Varespladib methyl

X
- Ximelagatran

Tamoxifen – little activity

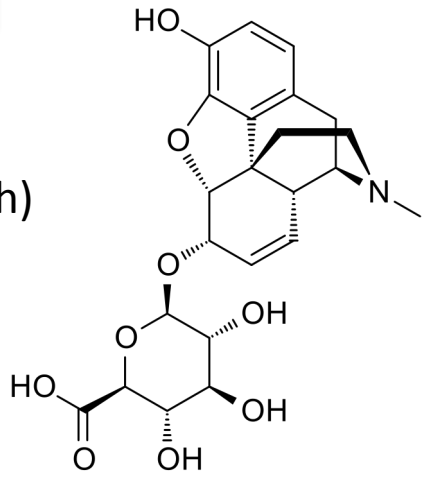


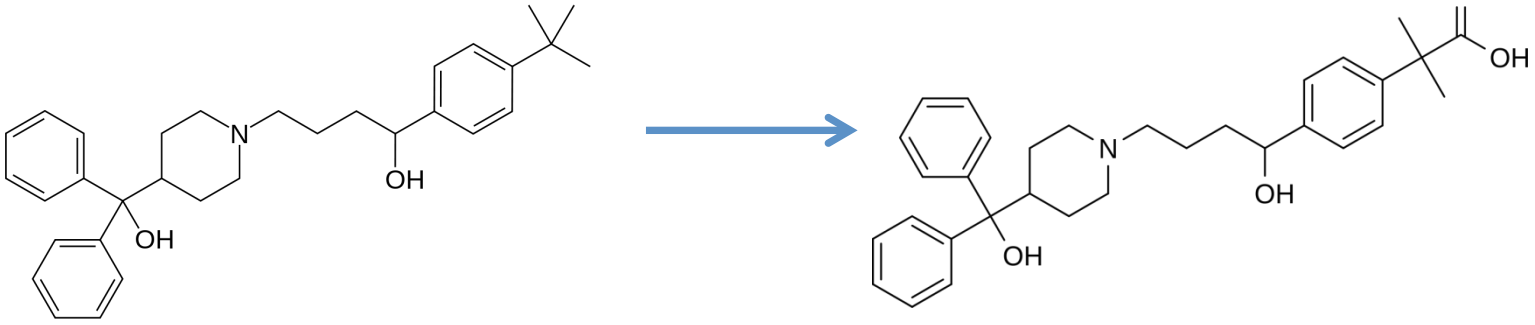


Afimoxifene – Active metabolite

The changing bioactivities of metabolites

Morphine is enzymatically activated to form sugar derivatives (morphine-glucuronides) that are (much) more active than the parent compound.

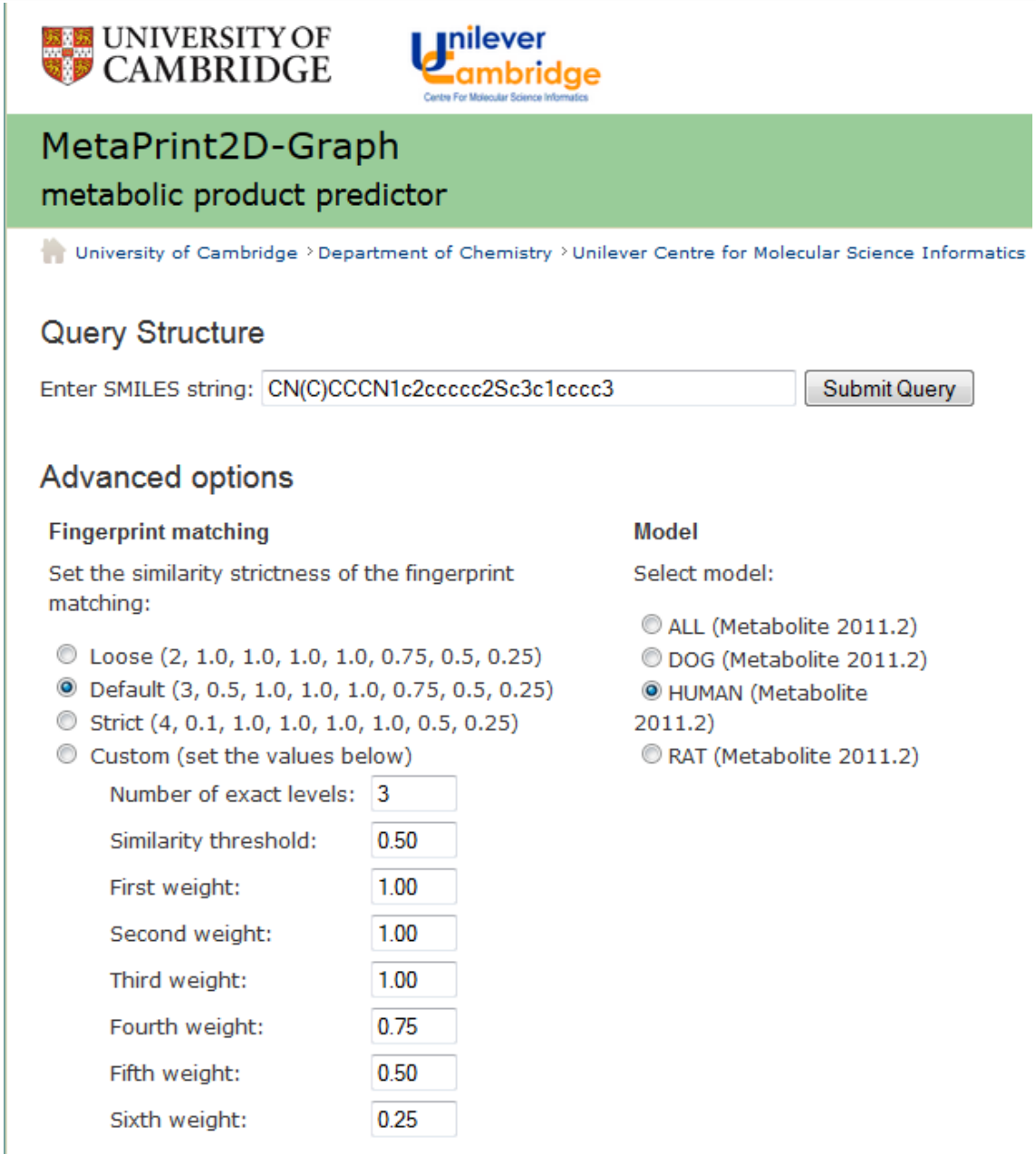




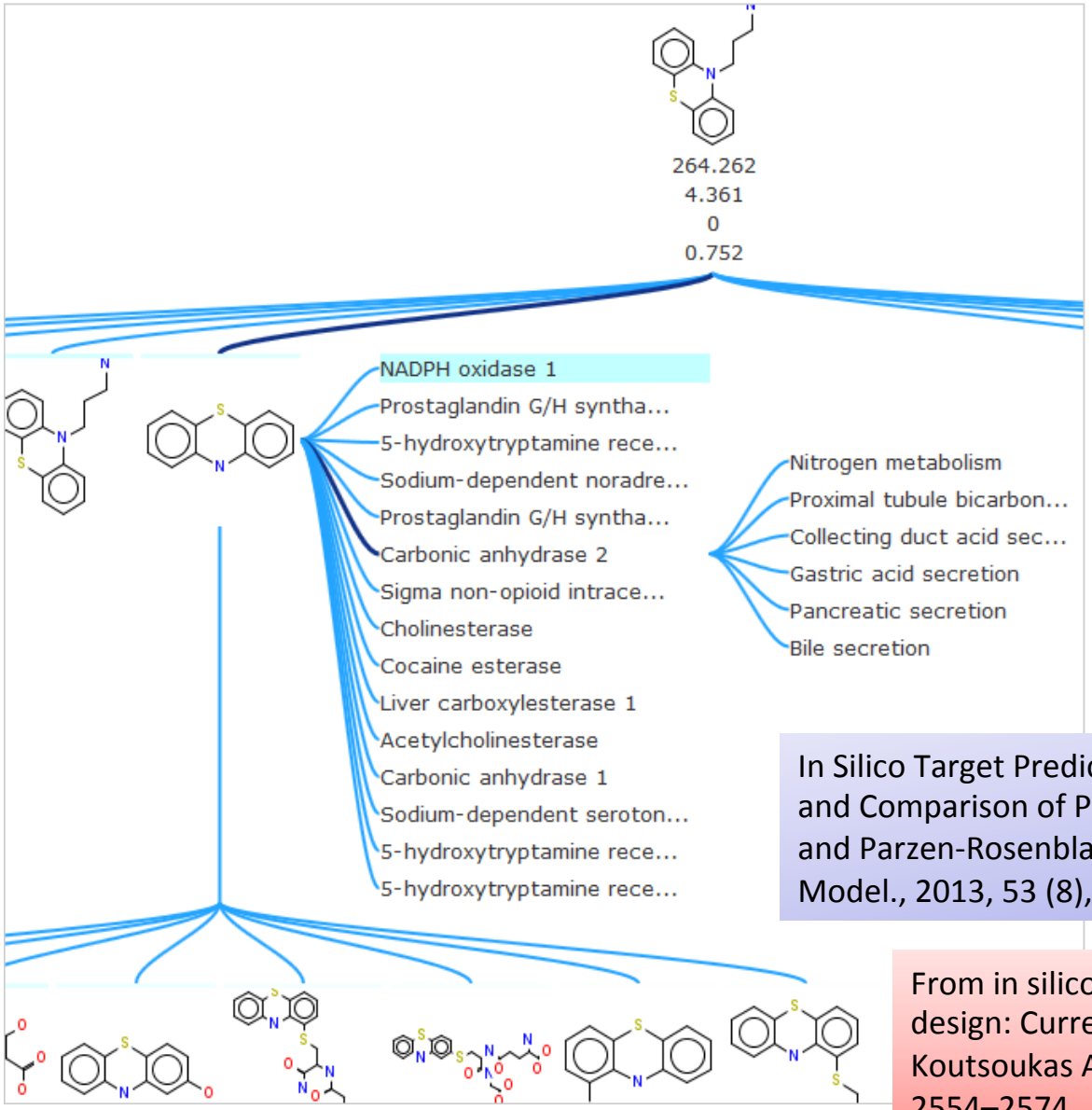
Terfenidine (pro-drug)
(Toxic if not metabolised, hERG)

Fexofenadine (active) – has to be Metabolised by the CYP3A4 isoform. If this is Missing, then very toxic (arrhythmia followed by cardiac arrest)

Example Promazine – predicted metabolites and activities



Predicted targets and Biological effects, which we may possibly relate to phenotypic changes (polypharmacology of parent AND metabolites)

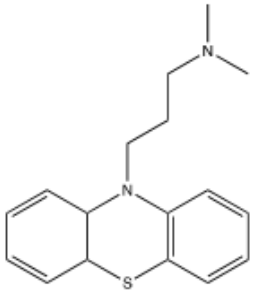


In Silico Target Predictions: Defining a Benchmarking Data Set and Comparison of Performance of the Multiclass Naïve Bayes and Parzen-Rosenblatt Window. Koutsoukas et al., J. Chem. Inf. Model., 2013, 53 (8), pp 1957–1966. DOI: 10.1021/ci300435j

From in silico target prediction to multi-target drug design: Current databases, methods and applications. Koutsoukas A., Journal of Proteomics, 2011, 74(12), 2554–2574

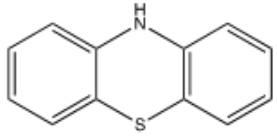
Polypharmacology and metabolism. Predicted and measured activities of promazine and metabolites

Predicted protein targets	Known binding affinities
HTR7	No ED
DRD3	Ki = 69 nM
CHRM4	Ki = 42 nM
DRD1	Ki = 1232 nM
HERG	No ED
CHRM5	Ki = 46 nM
OPRK1	No ED
HTR2C	Ki = 37 nM
HRH1	Ki = 0.4 nM
DRD4	No ED
SLC6A2	Ki = 13 nM
HRH2	No ED
ADRA1D	Ki = 3.74 nM
HTR6	No ED
CHRM3	Ki = 88 nM
HTR1A	No ED
SERT	Ki = 46 nM
SIGMAR1	Ki = 114 nM
CHRM1	Ki = 117 nM



A) Promazine
10-(3-dimethylaminopropyl)phenothiazine

N-Dealkylation



B) Phenothiazine core
Thiodiphenylamine

Predicted protein targets	Known binding affinities
Amine oxidase [flavin-containing] A	IC50 = 592 nM
Prostaglandin G/H synthase 1	IC50 = 196 nM
Sodium-dependent noradrenaline transporter	IC50 = 461 nM
Prostaglandin G/H synthase 2	IC50 = 532 nM

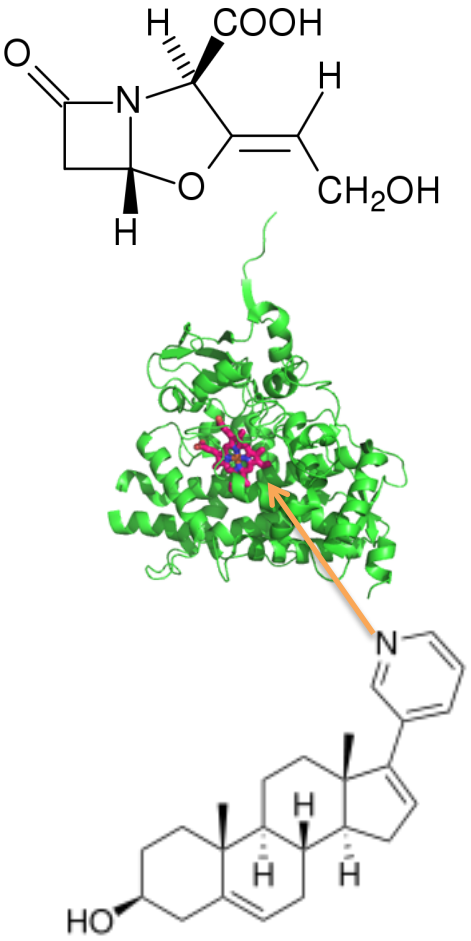
Promazine an antipsychotic agent (for schizophrenia) is an antagonist of dopamine receptors (DRD 1,2,3,4), muscarinic receptors (CHRM 1-5) and histamine receptor 1 (HRH1).
Predictions obtained for the parent (left).
Promazine is metabolized into the terminal metabolite phenothiazine core (thiodiphenylamine) by *N-dealkylation.*

The terminal metabolite was predicted to be active against Amine Oxidase, Cycloxygenase 1 and 2 (PTGS1 and PTGS2) and the Sodium-dependent noradrenaline transporter.

Blocking metabolism to promote efficacy

- Clavulanic acid is co-administered with amoxicillin to block β-lactamase. Clavulanic acid is a suicide inhibitor, covalently bonding to a serine residue in the active site of the β-Lactamase

- CYP17 is the crucial enzyme catalysing the conversion of pregnenolone and progesterone to dehydroepiandrosterone (DHEA) and androstenedione (which promote tumour growth) in gonadal and adrenal glands. Thus, blockade of androgen production in testes and adrenals by CYP17 (and in this case also CYP11B1) inhibition is being investigated for the treatment of prostate cancer.
- Nature Reviews Urology 11, 32–42 (2014) doi:10.1038/nrurol.2013.274



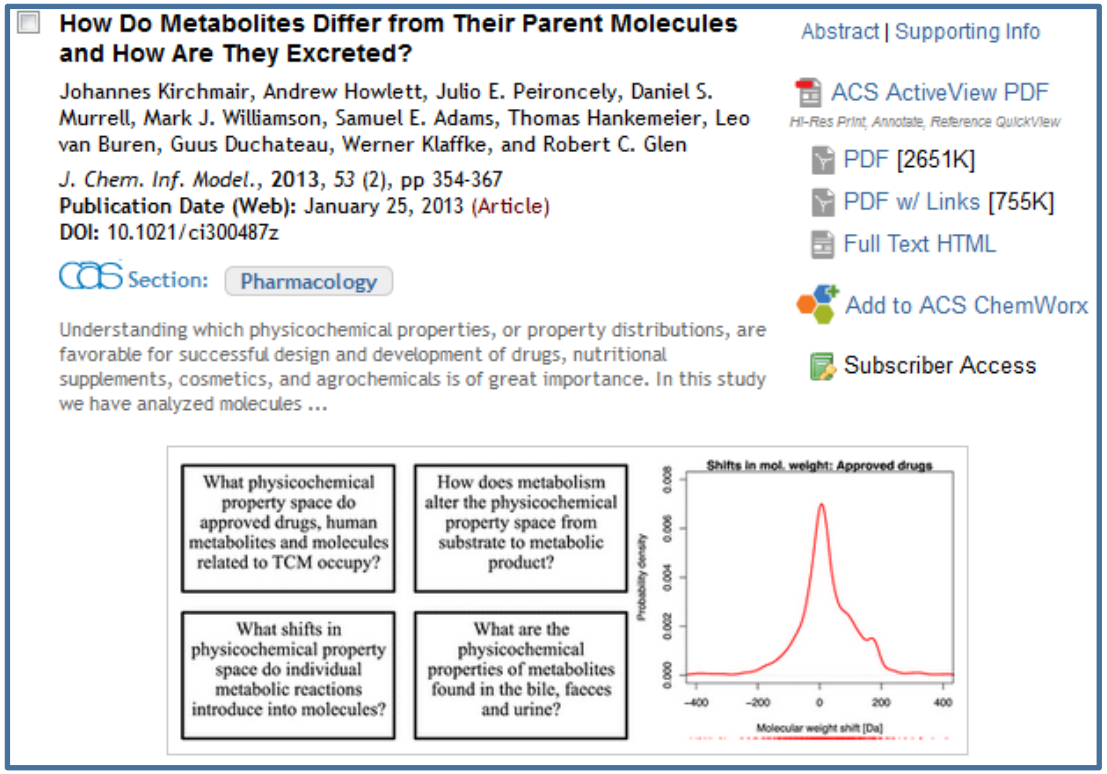
Abiraterone – a Cyp17-A1inhibitor

metabolism is very important for excretion of waste products

Drugs are eliminated from the body either unchanged as the parent drug or as metabolites.
Organs that excrete drugs eliminate polar compounds (water soluble) more readily than components with high lipid (fat) solubility. An exception is the lungs.
Lipid soluble drugs are not readily eliminated until they are metabolized to more polar compounds.

Possible sources of excretion include:

- Breath, Urine, Saliva
- Perspiration, Faeces
- Milk, Bile
- Hair, Skin



A “good” drug retires gracefully when it has completed it’s task

I would like to thank my collaborators and funders:

Acknowledgements

Mark Williamson
Rubben Torella
Sam Adams
Jon Tyzak
Johannes Kirchmair
Lora Mak
David Markus
Aixia Yan
Scott Boyer
Andreas Bender
Alison Choy

And Unilever,
AstraZeneca, BBSRC, MRC, Lhasa

UNIVERSITY OF CAMBRIDGE

Imperial College London