Predicting Routes, Sites and Products of Drug Metabolism

12th International ISSX Meeting, July 30th 2019

Matthew Segall – matt@optibrium.com

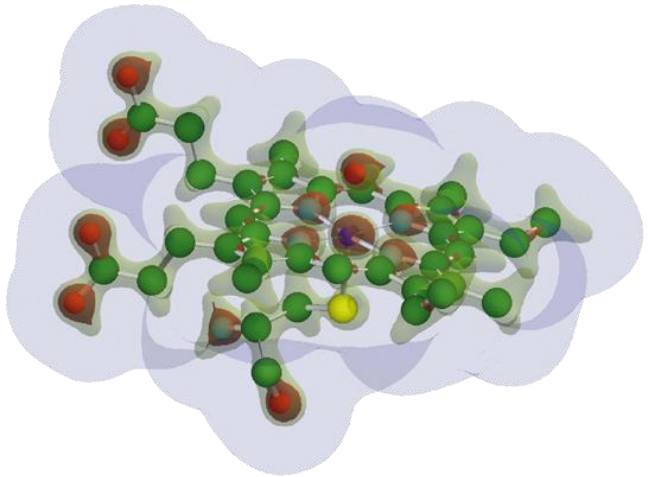
Overview

- Approaches to predicting metabolism
  - Empirical vs mechanistic
- Predicting P450 metabolism
  - P450 regioselectivity
  - WhichP450
- Beyond P450s
  - Flavin-containing monooxygenases (FMO)
  - UDP glucuronosyltrasfreases (UGT)
- Conclusions

Approaches to Predicting Drug Metabolism

Empirical

Mechanistic



Statistical Modelling

Machine Learning

Molecular Dynamics

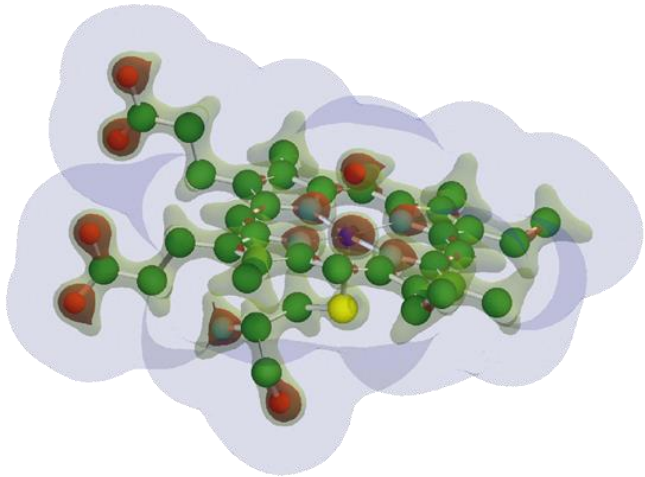
Quantum Mechanics

Approaches to Predicting Drug Metabolism

Empirical

Pros	Cons
Fast	Needs lots of data
Easy to set up	Non-transferable
	Qualitative

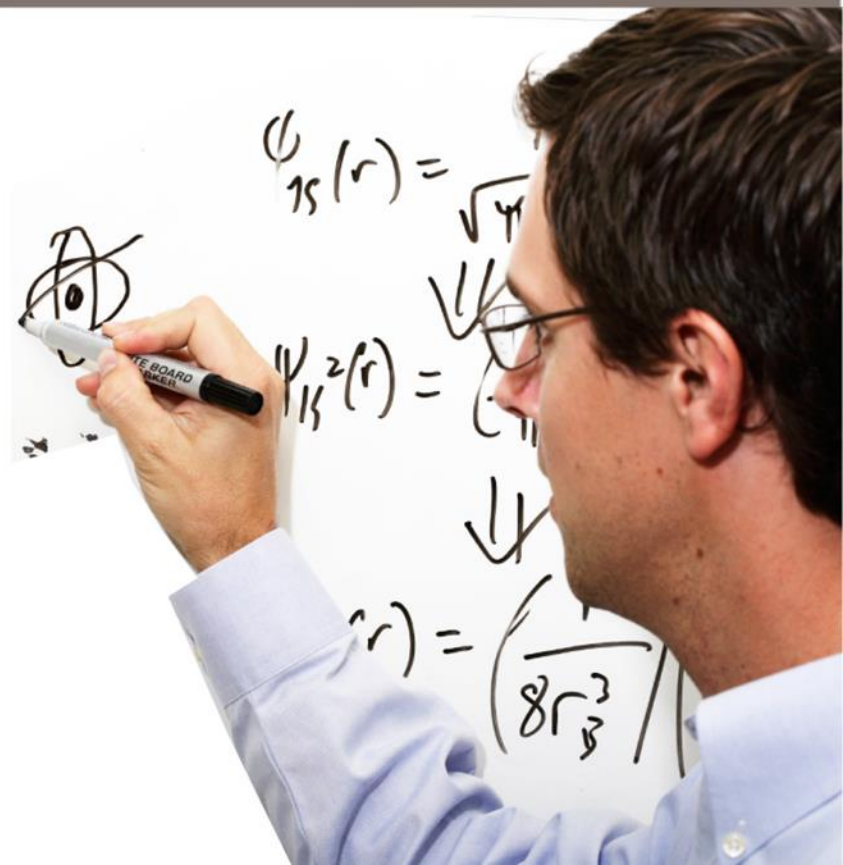
Mechanistic



Pros	Cons
Can be built on smaller high-quality data sets	(Very) slow
Transferable – based on physical principles	Requires detailed understanding
(Semi) quantitative	

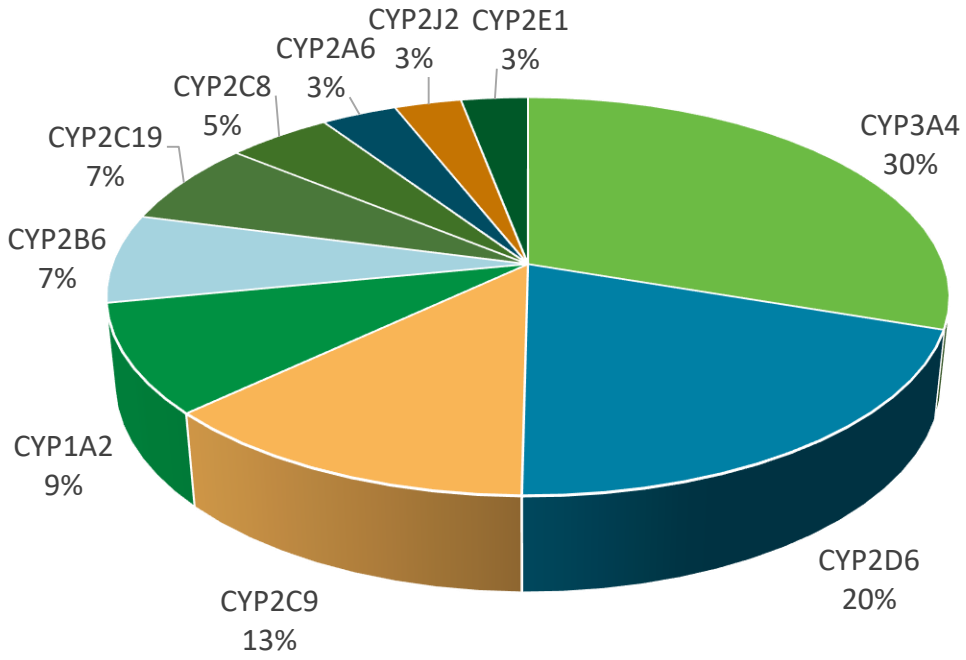
Predicting P450 Metabolism

optibrium



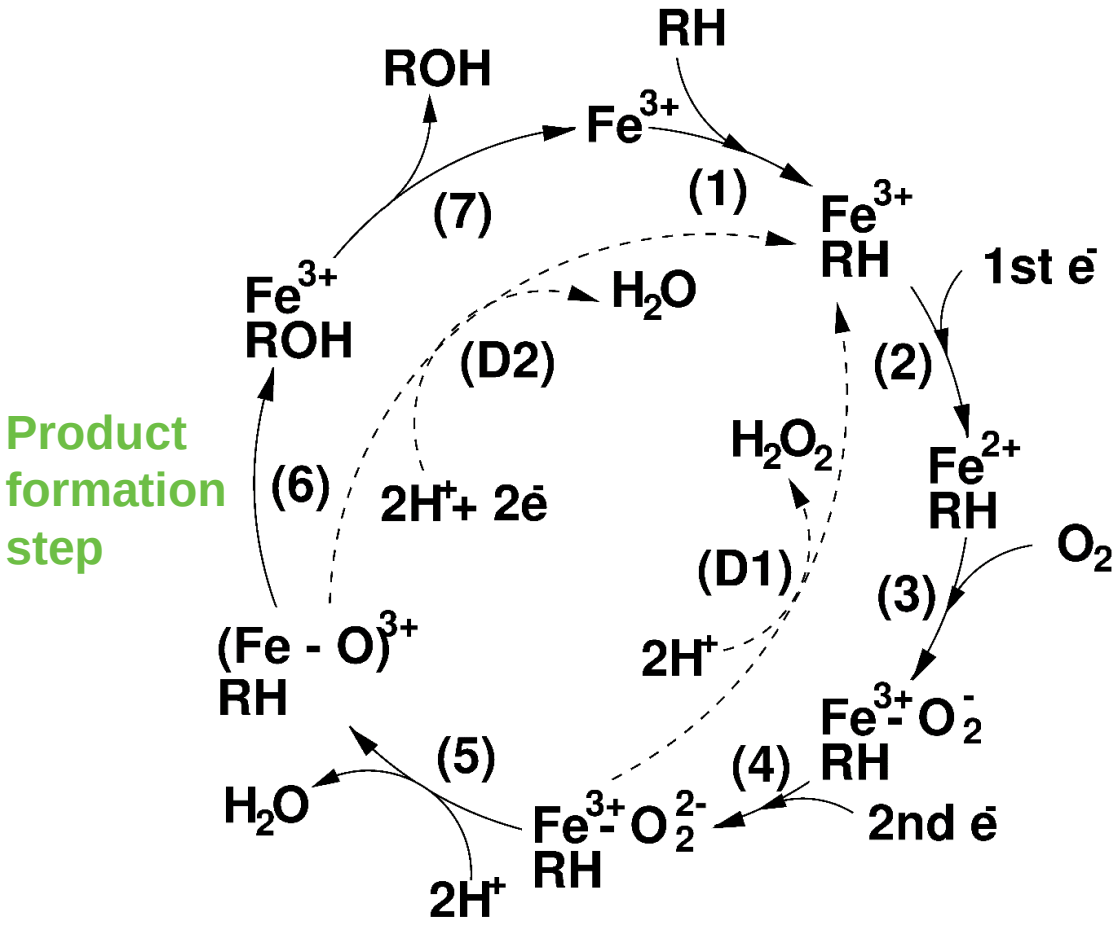
Cytochrome P450s

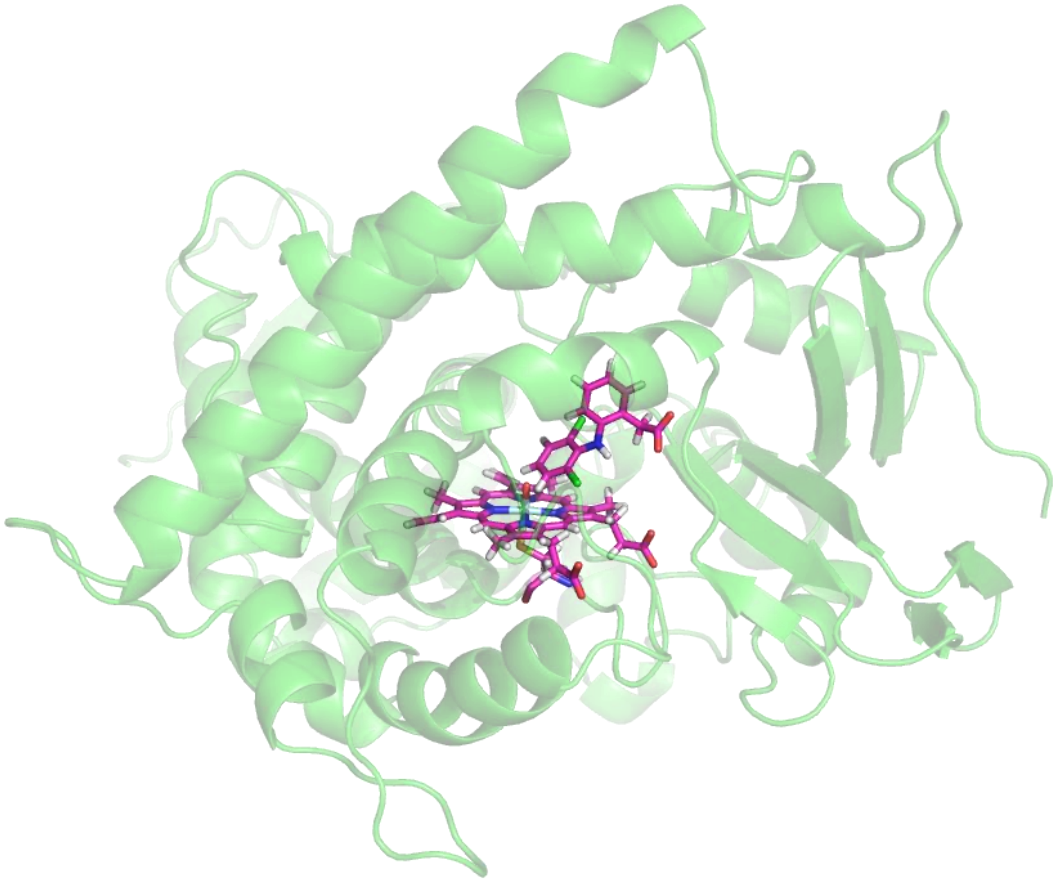
- Ubiquitous superfamily of haem-containing monooxygenase enzymes
- Responsible for ~70-80% of phase I drug metabolism, leading to:
  - Rapid clearance or low bioavailability
  - Potential for drug-drug interaction
  - Impact of P450 polymorphism
  - Bioactivation to form reactive/toxic metabolites
- Primary isoforms responsible for drug metabolism in humans



P450 Catalytic Cycle

Predicting sites of metabolism - regioselectivity





P450 compound I and bound substrate

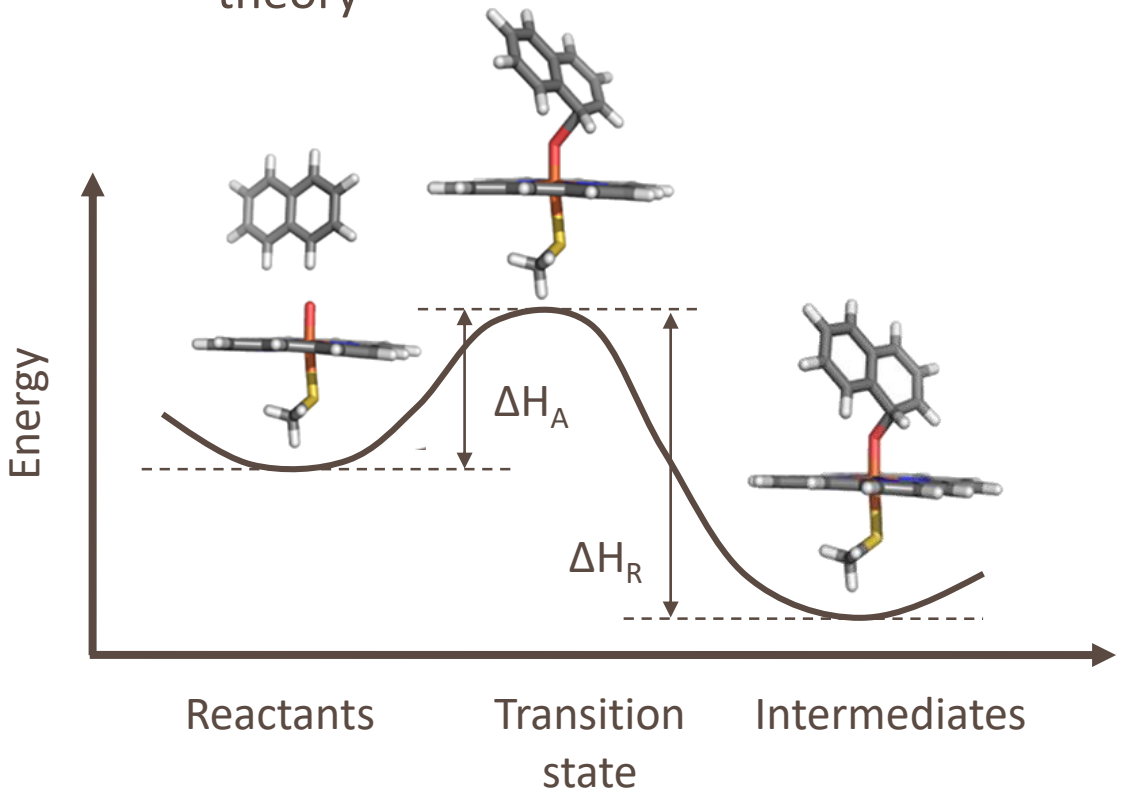
Predicting Sites of Metabolism
Regioselectivity

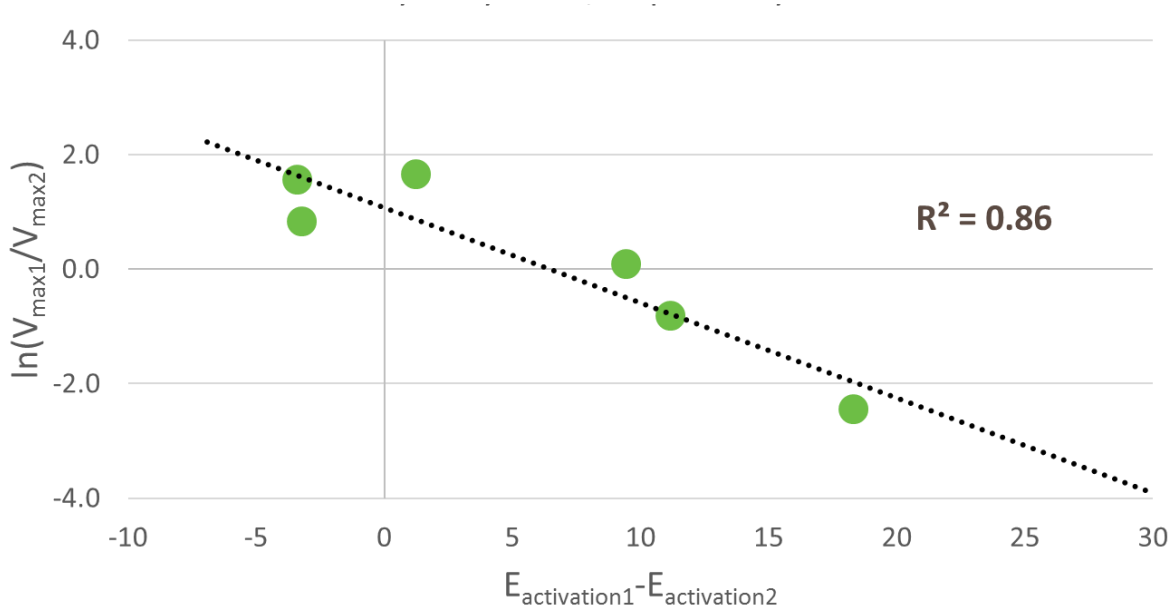
Two primary factors determine the sites of metabolism:

- Electronic properties of substrate – reactivity
  - *H*-abstraction – aliphatic oxidation, *N*-dealkylation, *O*-dealkylation
  - Direct oxidation – aromatic oxidation, epoxidation, *N*-oxidation, *S*-oxidation
  - Independent of isoform
- Orientation of substrate in active site
  - Electrostatic interactions between protein and substrate
  - Freedom to move
  - Steric accessibility
  - Dependent on isoform and substrate

Quantum Mechanical Models for CYP Reactivity

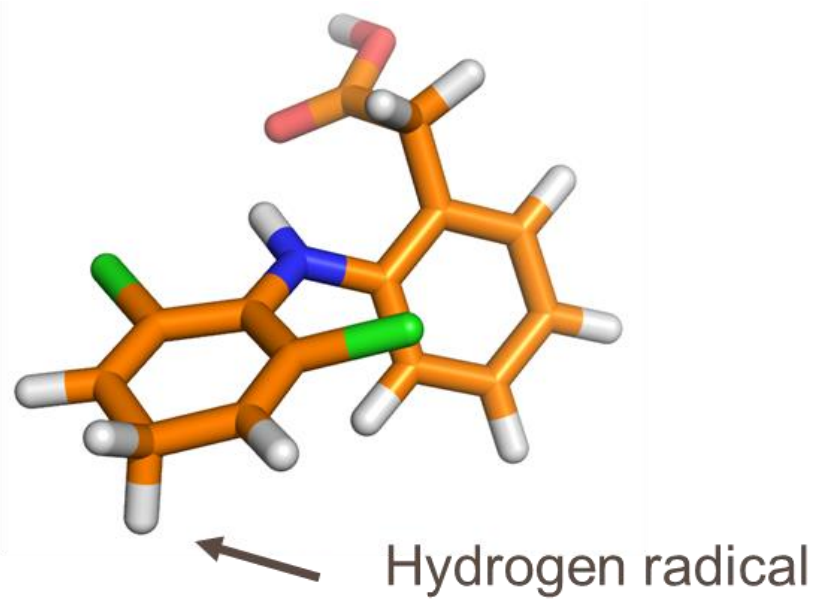
- The activation energy (ΔH_A) of the rate-limiting step is a key factor determining the rate of reaction at each site
  - Reaction energetics modelled for *H*-abstraction and direct oxidations using density functional theory





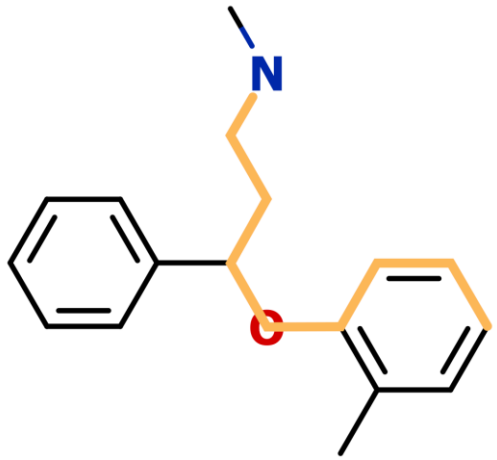
Quantum Mechanical Models for CYP Reactivity

- Semi-empirical QM methods (AM1) are used for practical calculations
  - Surrogate radical used instead of haem
  - Brønsted relationships used to estimate activation energies
  - Corrections applied based on *ab initio* QM
- Full substrate included in simulation
  - Not 'pattern matching' sites to precalculated energies
  - Includes subtle longer range effects
  - Important when developing a lead series



Capturing Steric and Orientation Effects

- Corrections to activation energies estimated for each isoform
  - 3A4, 2D6, 2C9, 1A2, 2C8, 2C19, 2E1
- Statistical models using 2D descriptors
  - Distances to charged functionalities, H-bond acceptors/donors, etc.
  - Distances to rings, flexible linkers, 'bulky' groups
- Trained and tested using high-quality regioselectivity data sets

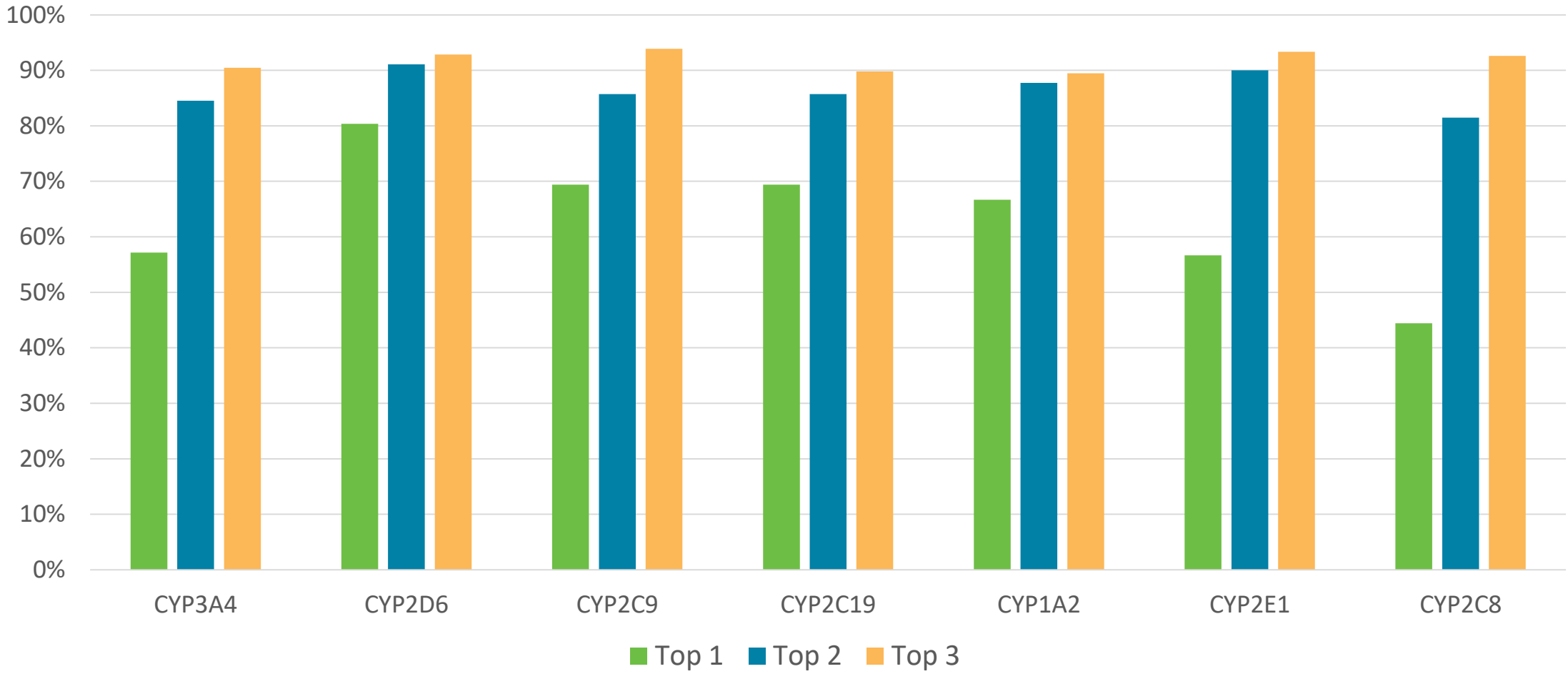


Isoform	Number of molecules
CYP3A4	305
CYP2D6	202
CYP2C9	193
CYP1A2	201
CYP2C19	184
CYP2E1	105
CYP2C8	106

J Tyzack *et al.* (2017) J. Chem. Inf. Model **56**(1) pp. 2180-2193

Validation

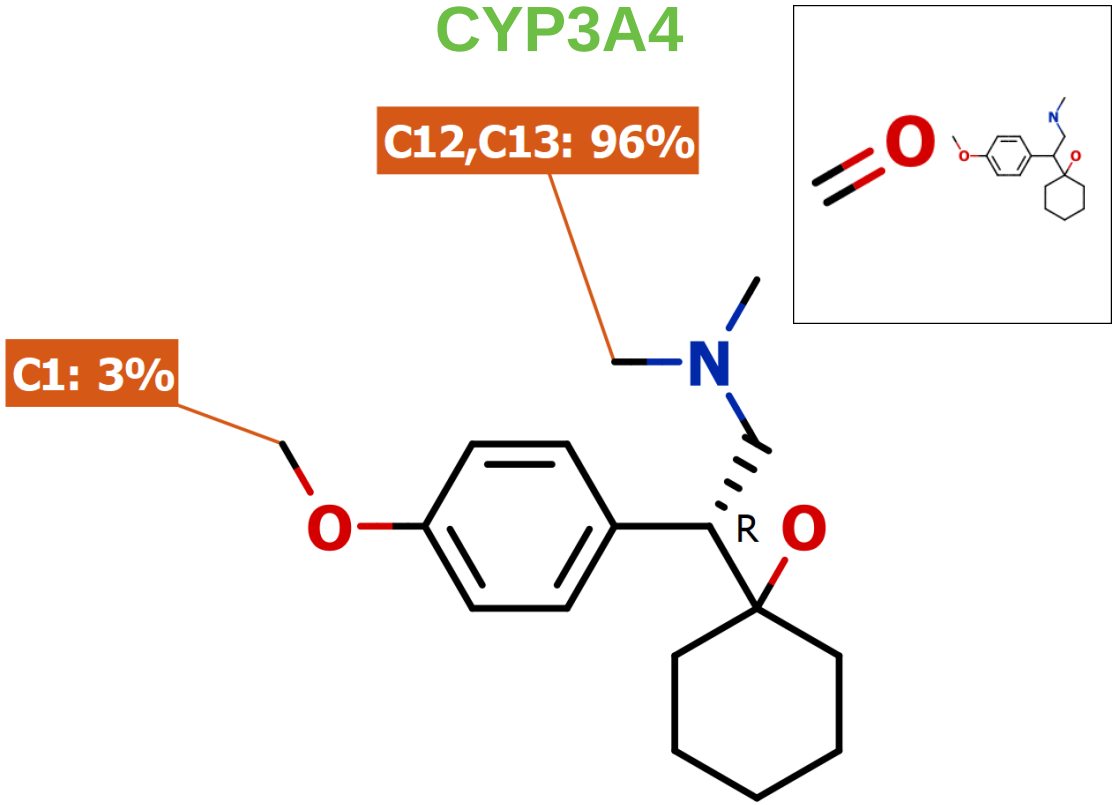
Independent test set of 30% of data



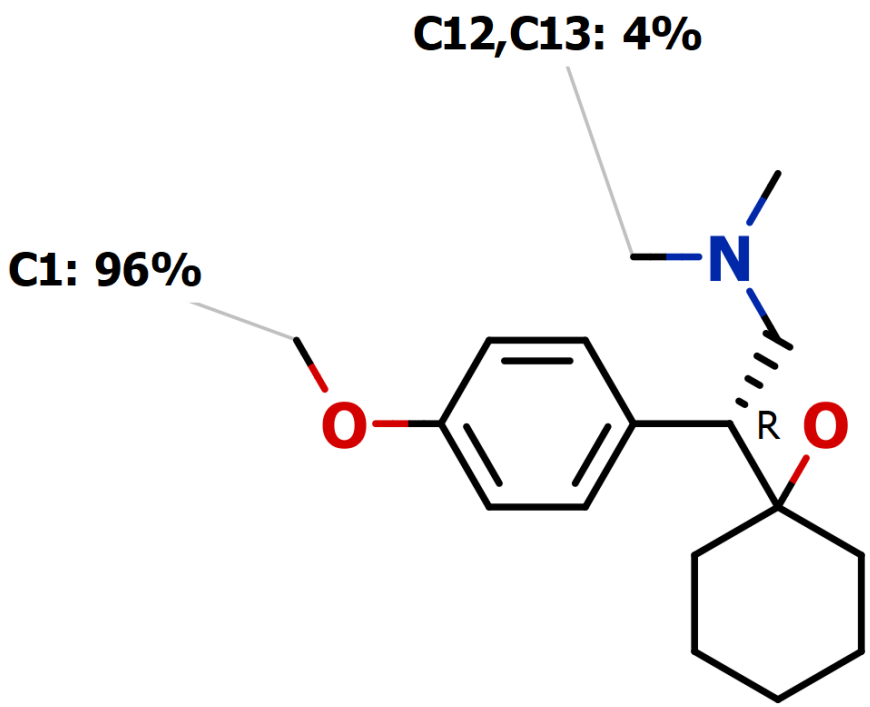
J Tyzack *et al.* (2017) J. Chem. Inf. Model **56**(1) pp. 2180-2193

Example Regioselectivity Prediction
Venlafaxine

CYP3A4

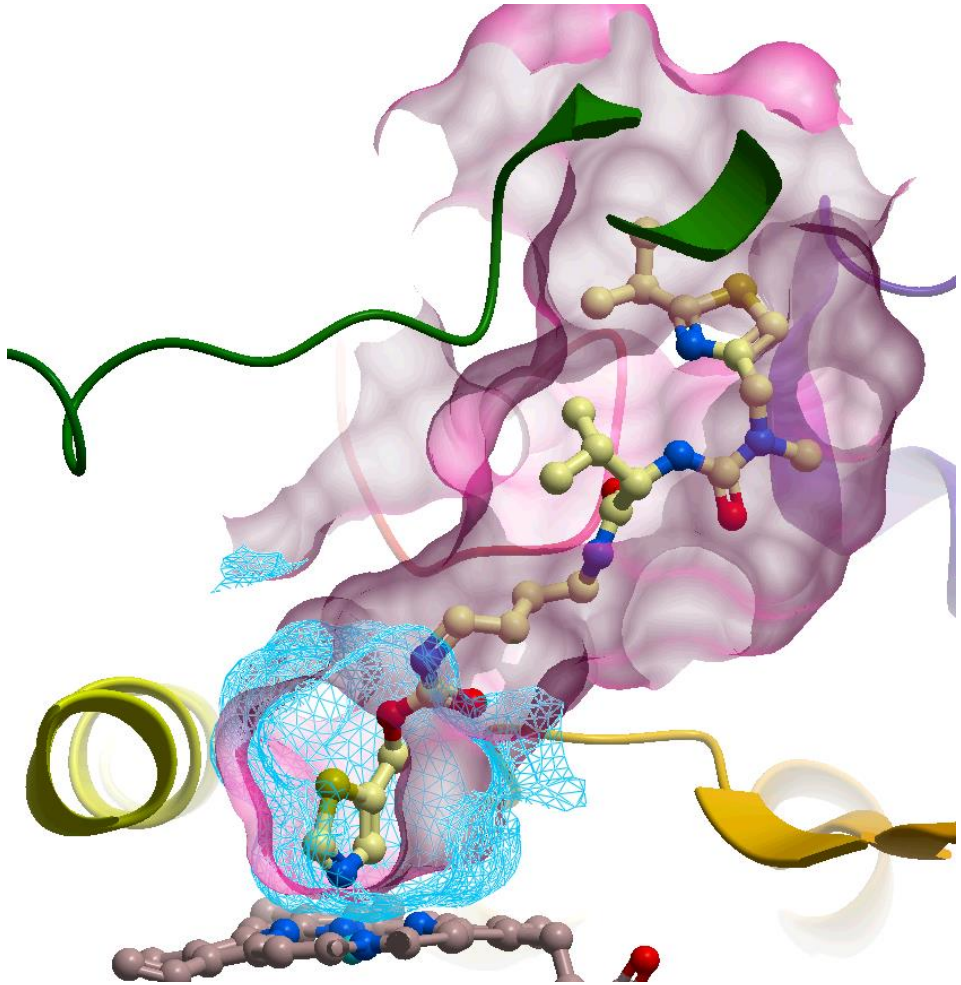


CYP2D6



WhichP450
Objectives

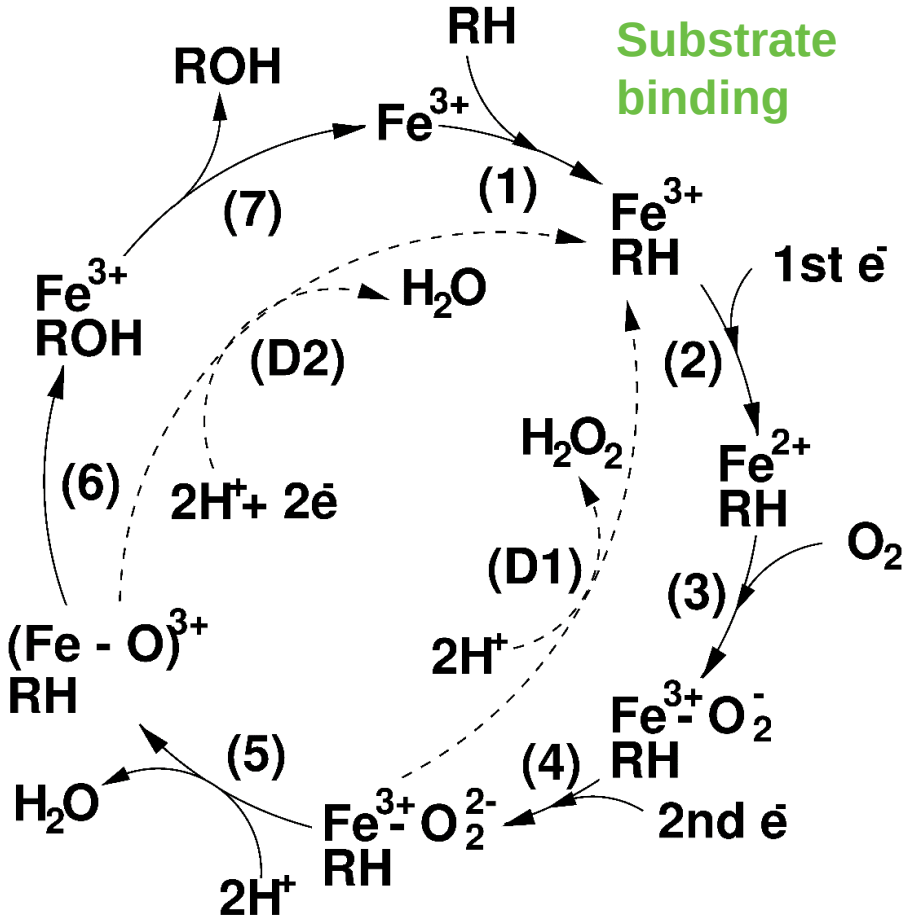
- Many isoforms of P450
  - Different active site constraints
- Predictions of regioselectivity for which isoform(s) are most relevant?
- Identify possibilities of DDIs or polymorphic effects
- Compounds may be metabolised by multiple isoforms

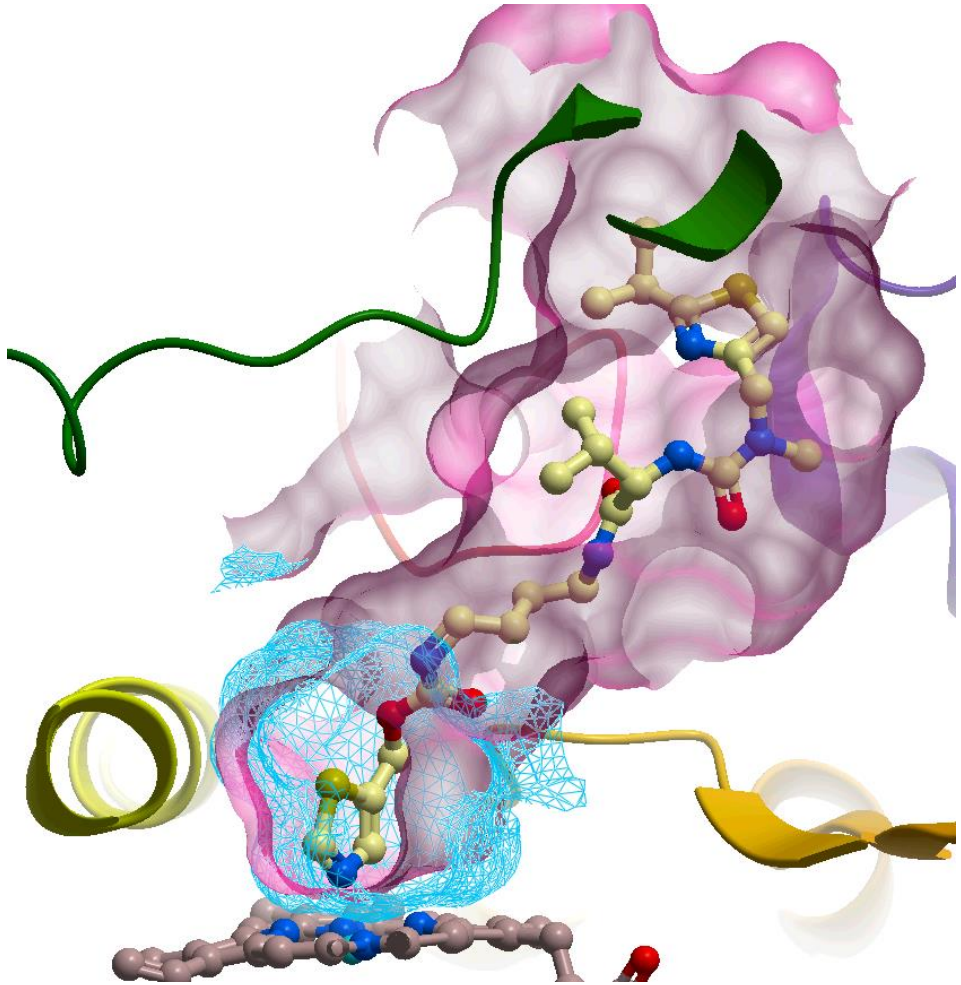


Binding sites: CYP3A4 – purple & CYP2E1 – blue

P450 Catalytic Cycle

Predicting which P450 isoform(s)

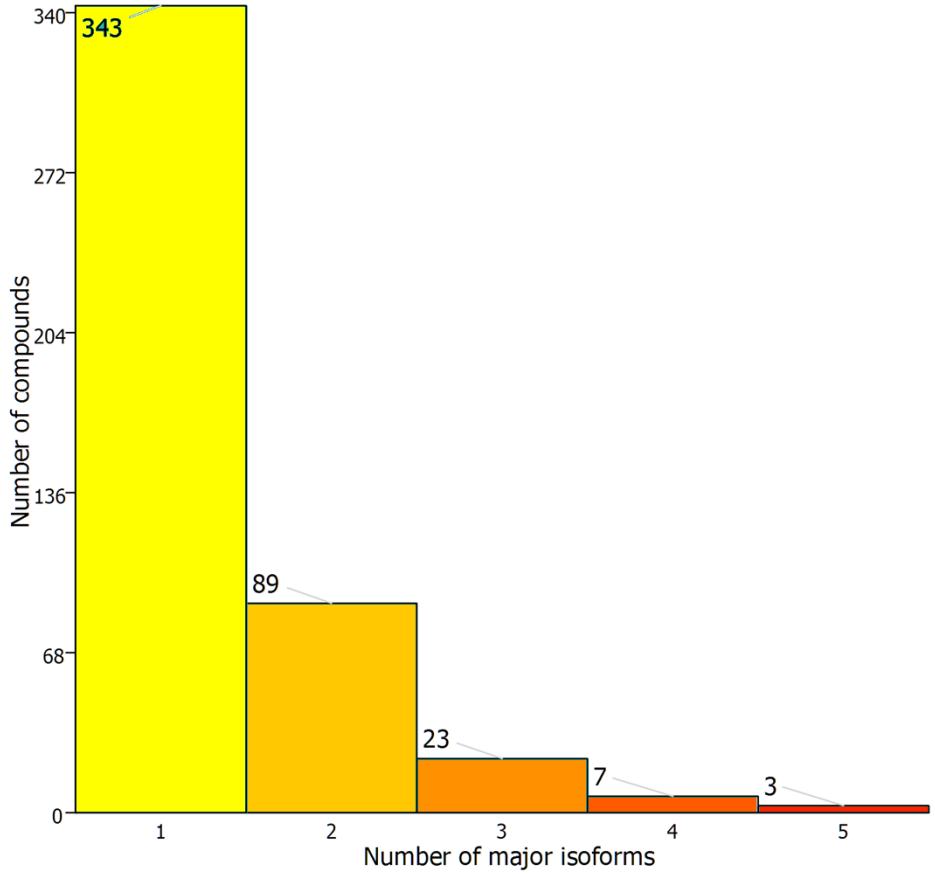




Binding sites: CYP3A4 – purple & CYP2E1 – blue

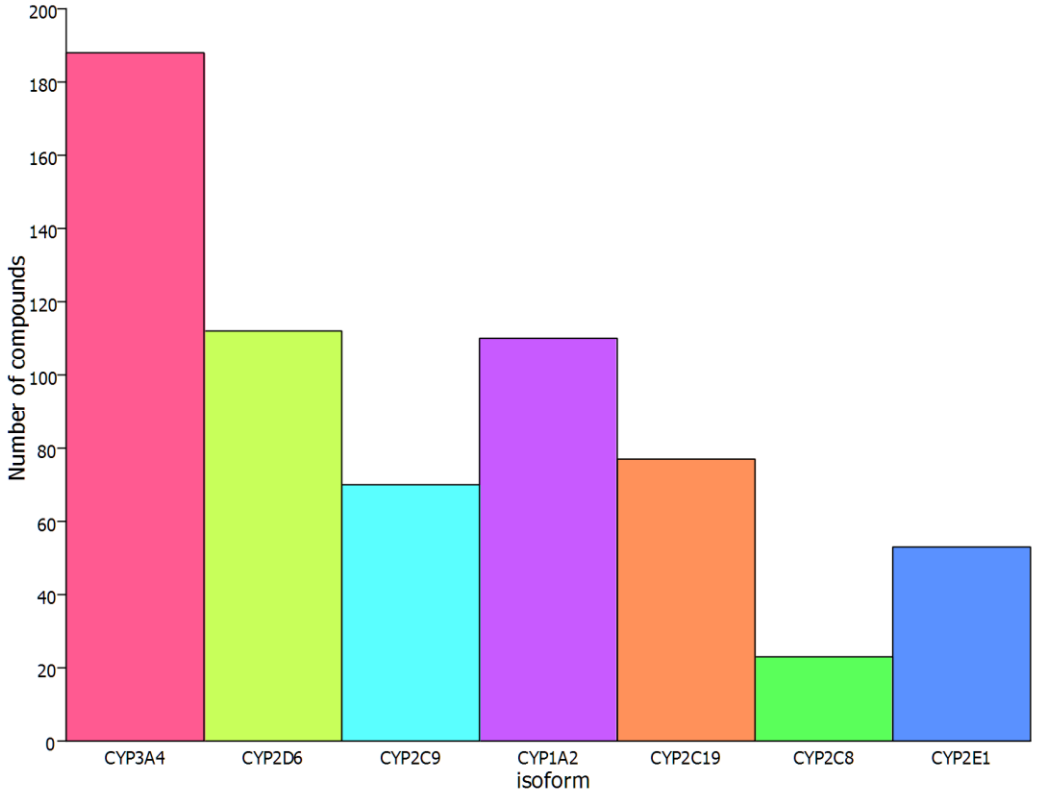
WhichP450
Methods

- Data set
  - 465 unique compounds
  - 633 compound/isoform pairs
- Considers 7 isoforms
  - 3A4, 2D6, 2C9, 1A2, 2C8, 2C19, 2E1
- Random forest model
  - Random forests
  - Whole molecule and 2D descriptors
- Model rank orders isoforms by probability

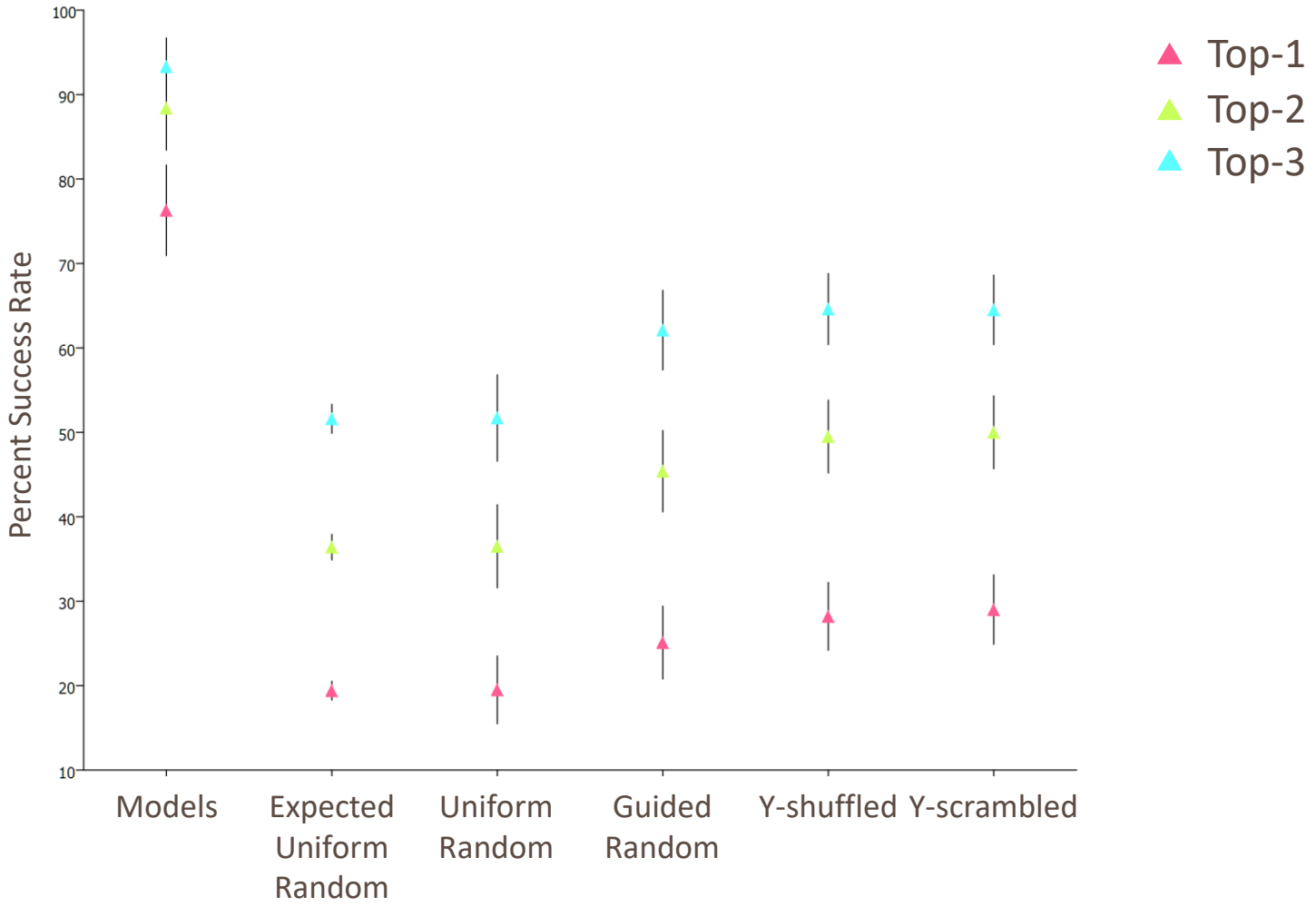


WhichP450
Methods

- Data set
  - 465 unique compounds
  - 633 compound/isoform pairs
- Considers 7 isoforms
  - 3A4, 2D6, 2C9, 1A2, 2C8, 2C19, 2E1
- Random forest model
  - Random forests
  - Whole molecule and 2D descriptors
- Model rank orders isoforms by probability

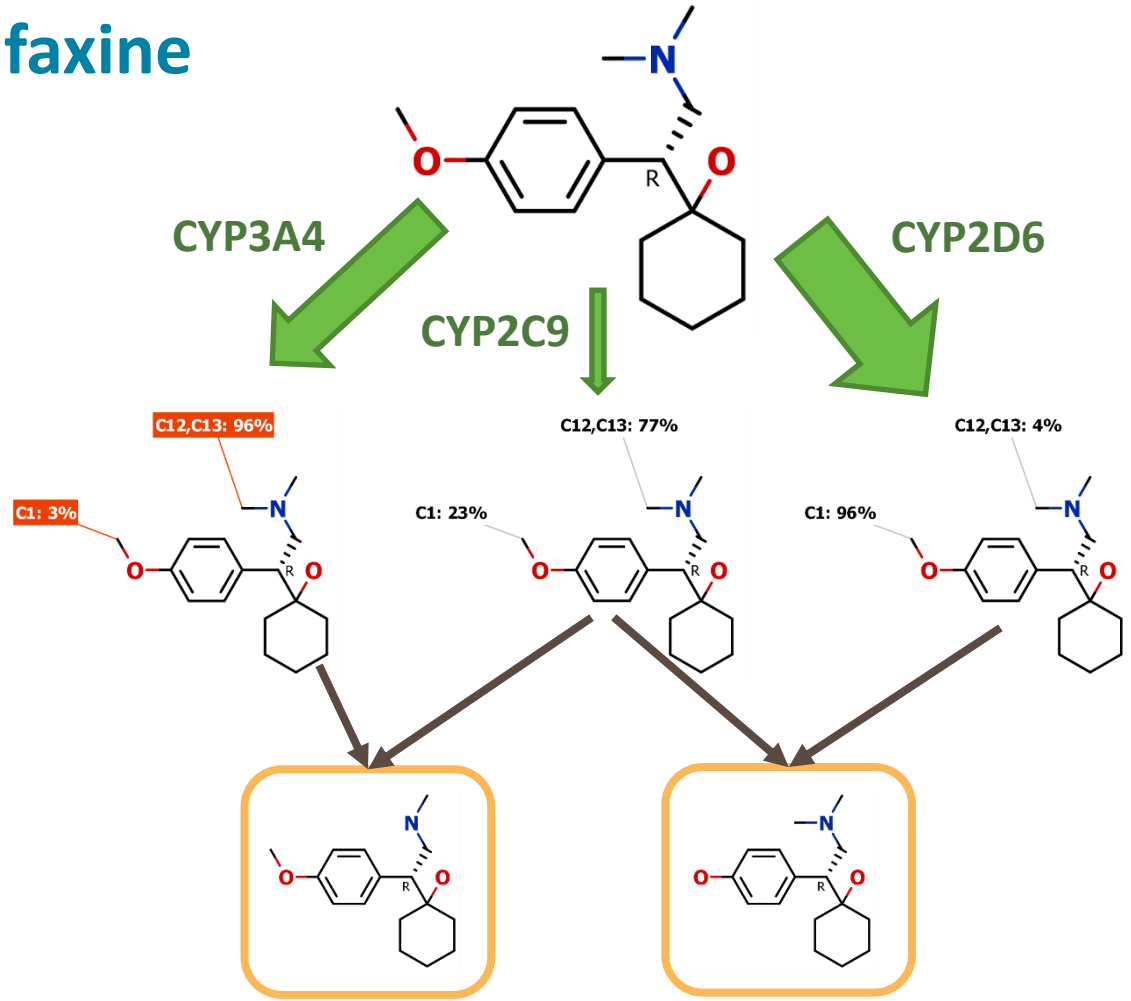


WhichP450
Results – Top-*k*



Putting it Together

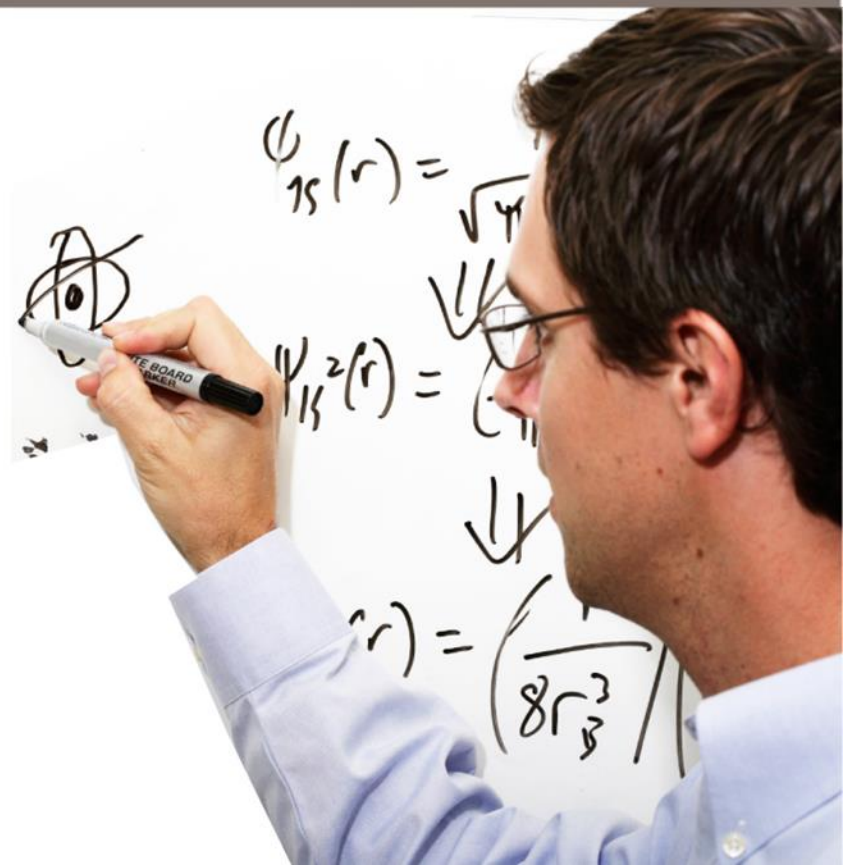
Venlafaxine



2C19 is also a minor isoform, but not predicted

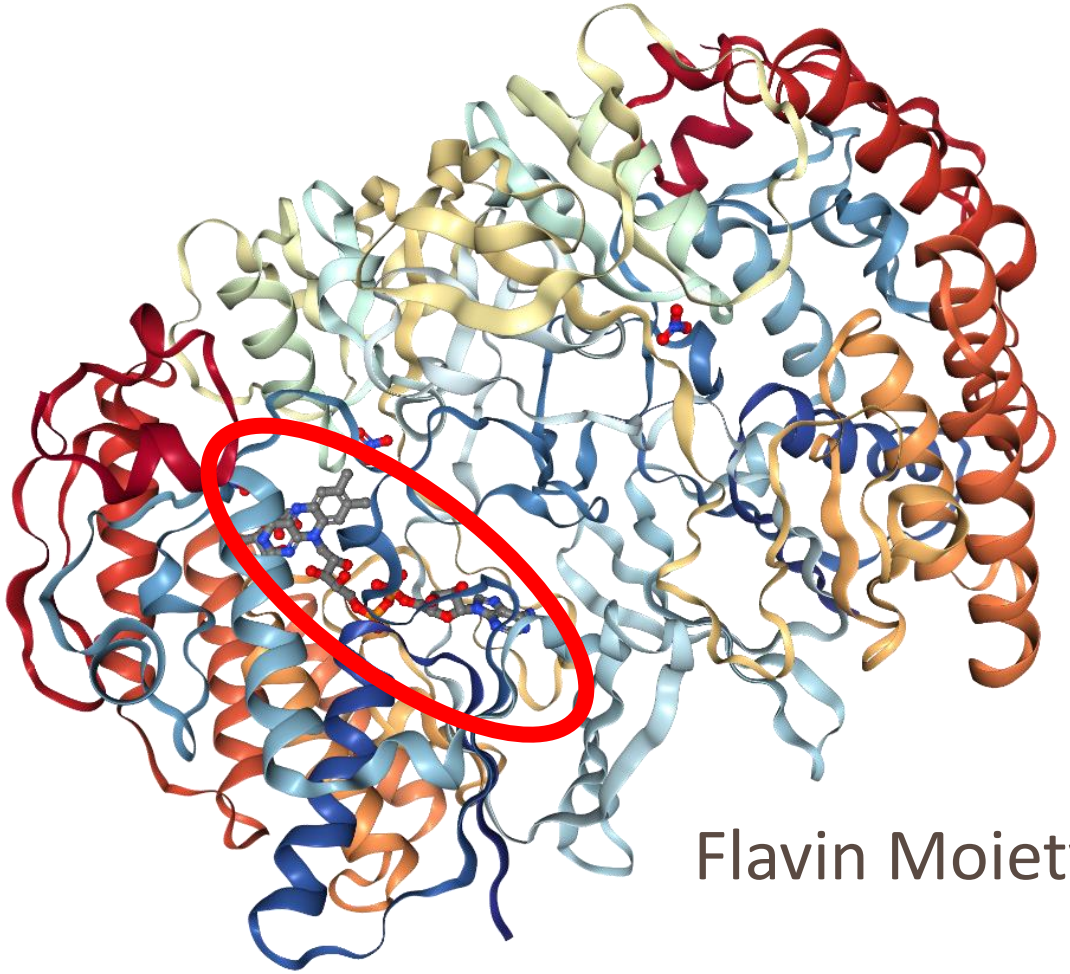
Beyond P450s

optibrium



Flavin-containing Monooxygenase (FMO)

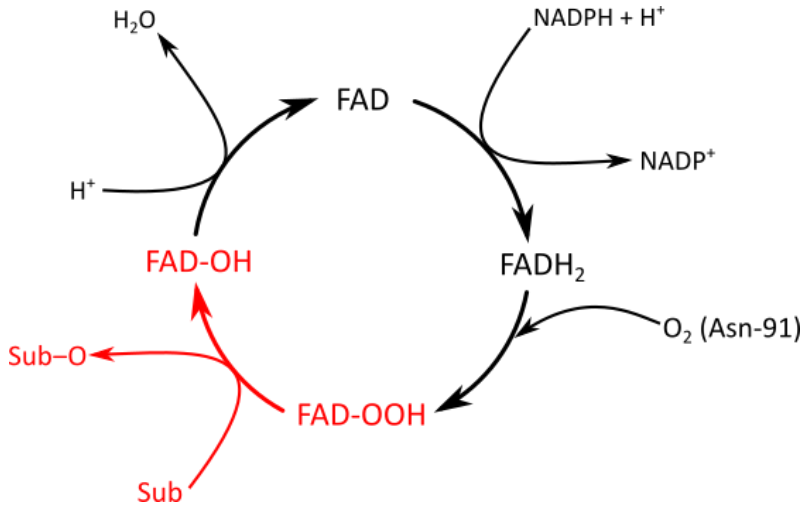
- Phase I enzyme class involved in compound metabolism
  - Found in multiple tissues
- 5 active isoforms (FMO1–5)
  - FMO3 major isoform found in adult liver
- Mechanism involves transfer of Oxygen from FAD–OOH
  - Predominantly *N*/*S*-oxidation

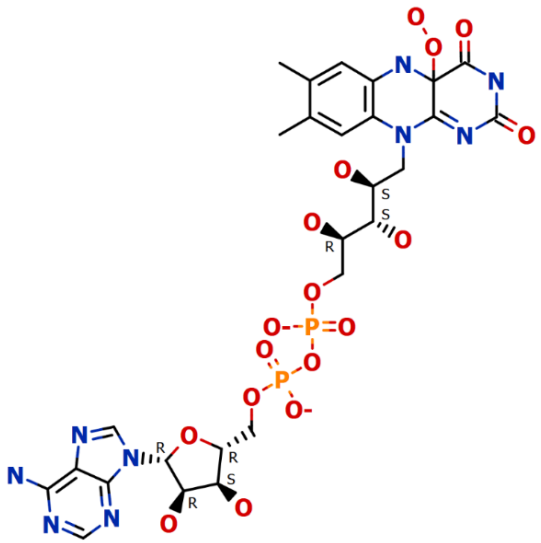


PDB 2XVE

Flavin-containing Monooxygenase (FMO)

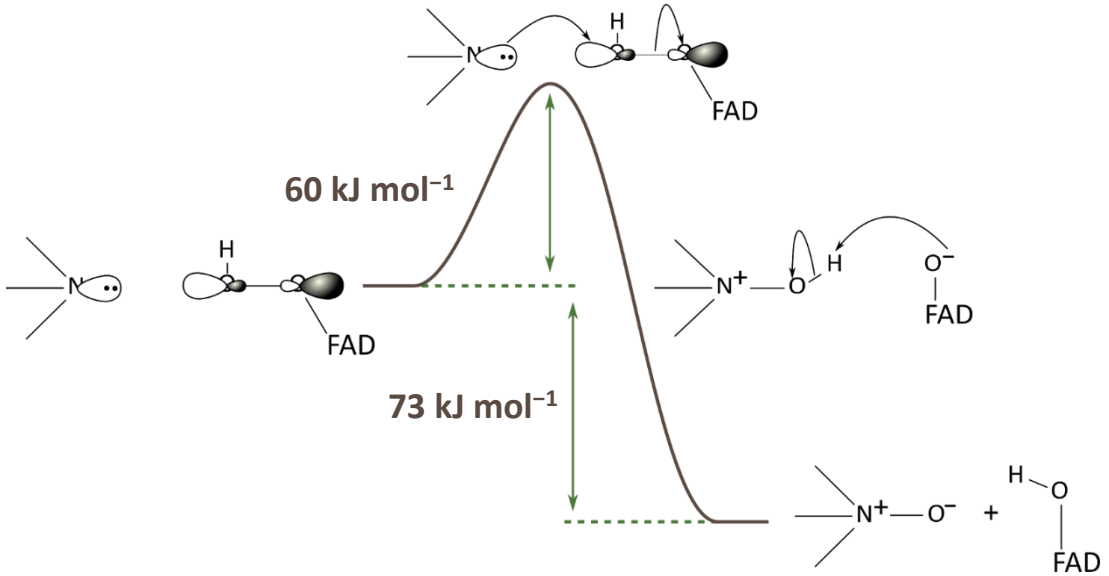
- Phase I enzyme class involved in compound metabolism
  - Found in multiple tissues
- 5 active isoforms (FMO1–5)
  - FMO3 major isoform found in adult liver
- Mechanism involves transfer of Oxygen from FAD–OOH
  - Predominantly *N*/*S*-oxidation

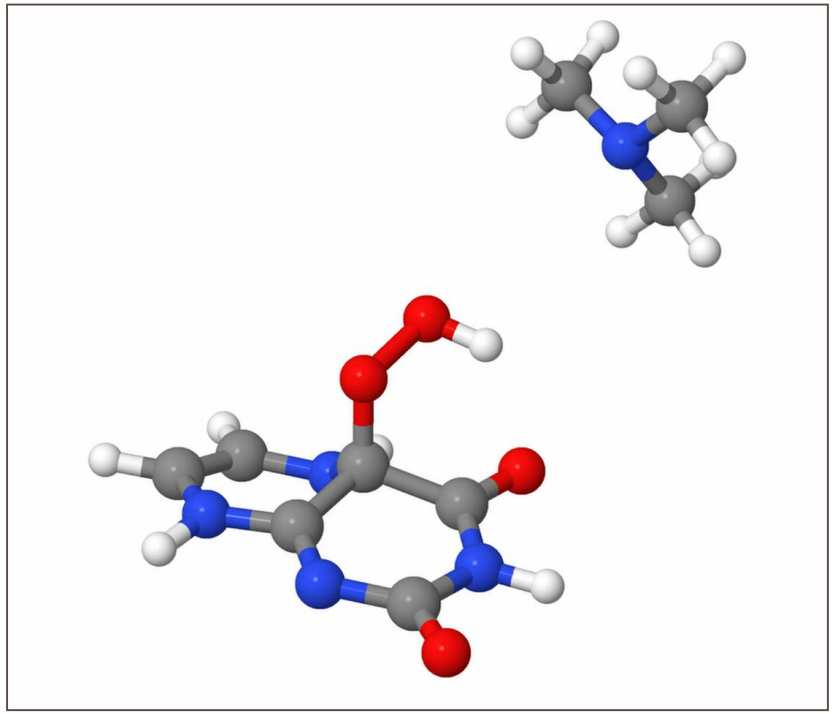




Modelling the Reaction Mechanism

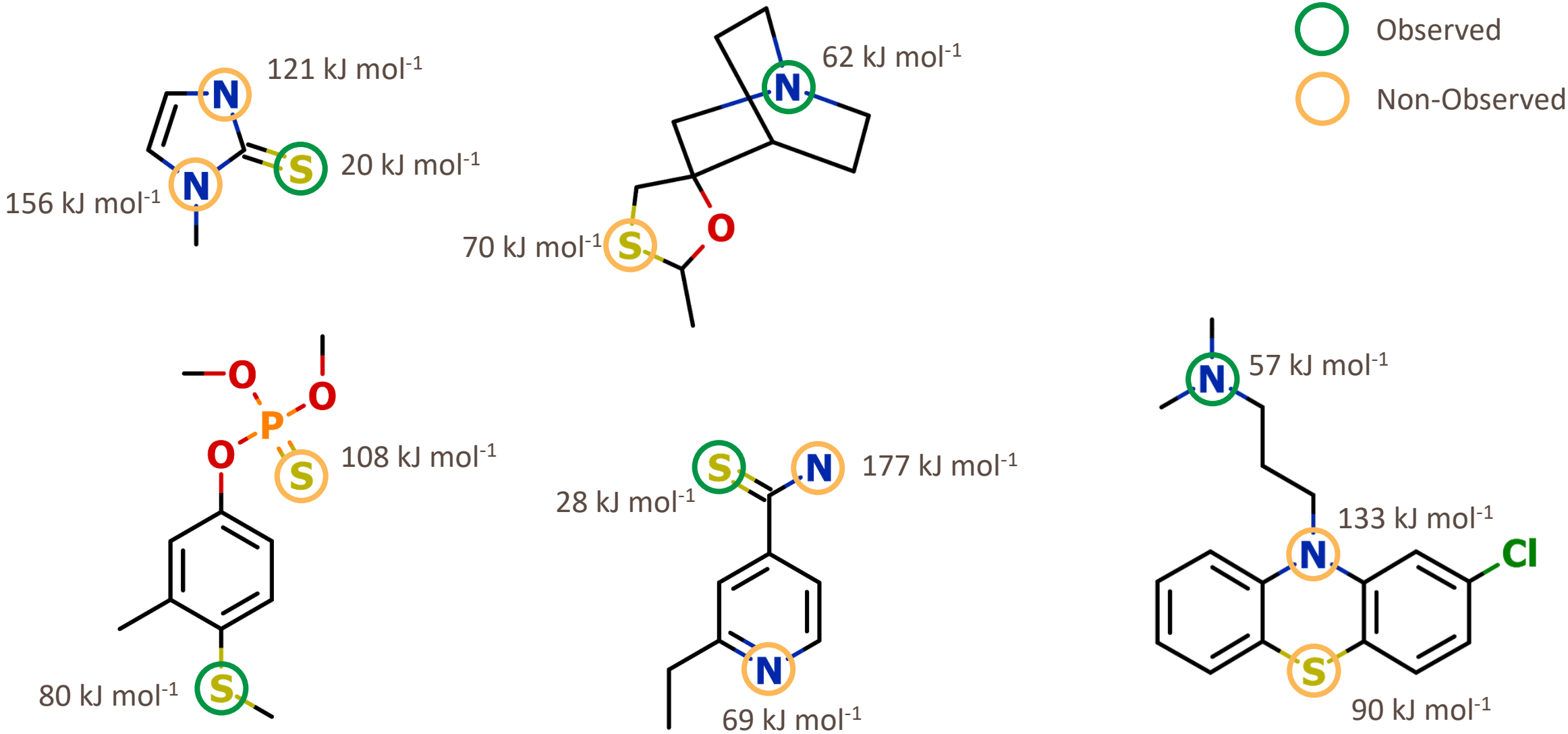
- QM simulations using DFT to determine reaction mechanism
  - Concerted, S_N2
- Calculate activation energy, ΔH_A





Identifying of Sites of FMO Metabolism

Activation Energies

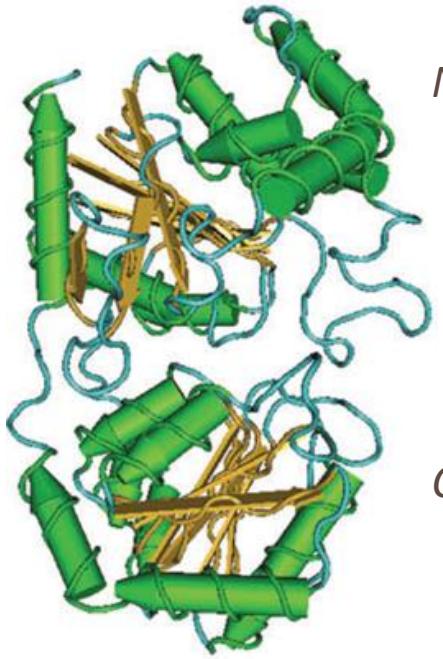


Example – Predicting FMO3 Metabolism

- Activation energies calculated with semi-empirical QM model of transition state
- Steric and orientation descriptors included
- Data set
  - 67 molecules
  - 210 potential sites of metabolism
- Gaussian processes machine learning
- Classification of potential sites as metabolised (True) or not (False)
- Results on independent test set
  - Kappa = 0.82
  - Accuracy 92%

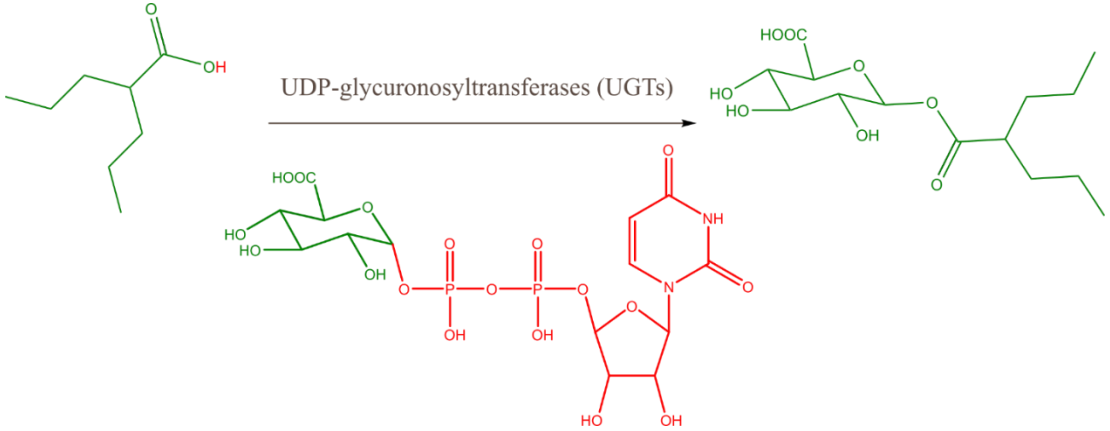
UDP-Glucuronosyltransferase (UGT)

- Major contributors to phase II metabolism
  - ~40% of all conjugation reactions
- Conjugation of substrate with glucuronic acid
- Several human isoforms implicated in drug metabolism
  - UGT1A – 1A1, 1A4, 1A9
  - UGT2B – 2B4, 2B7, 2B15



N-terminal domain
Substrate binding

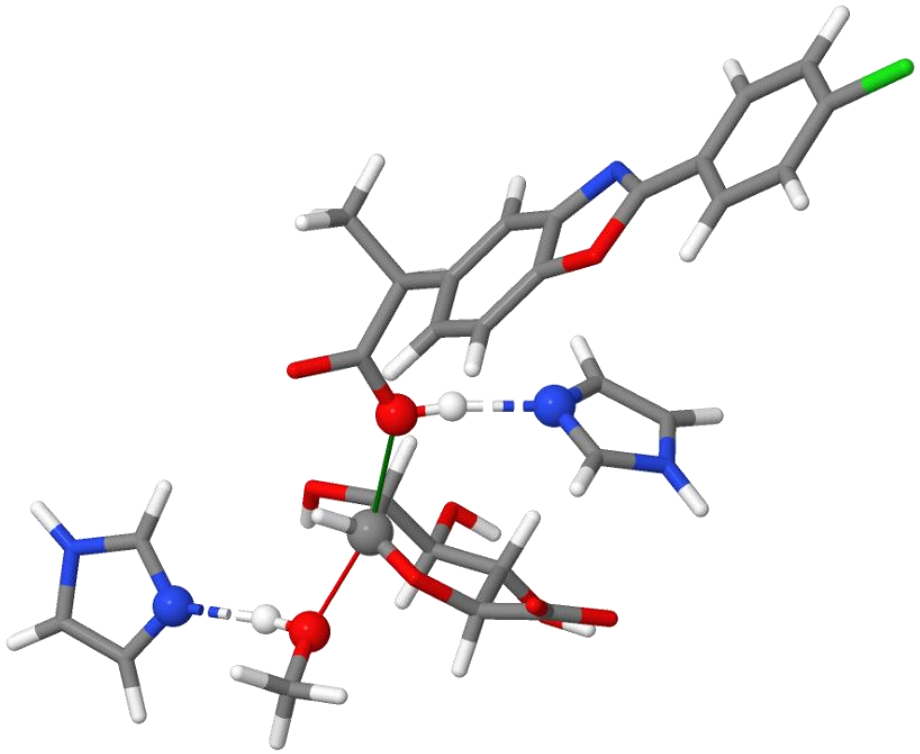
C-terminal domain
UDP-GA binding



PDB 2O6L

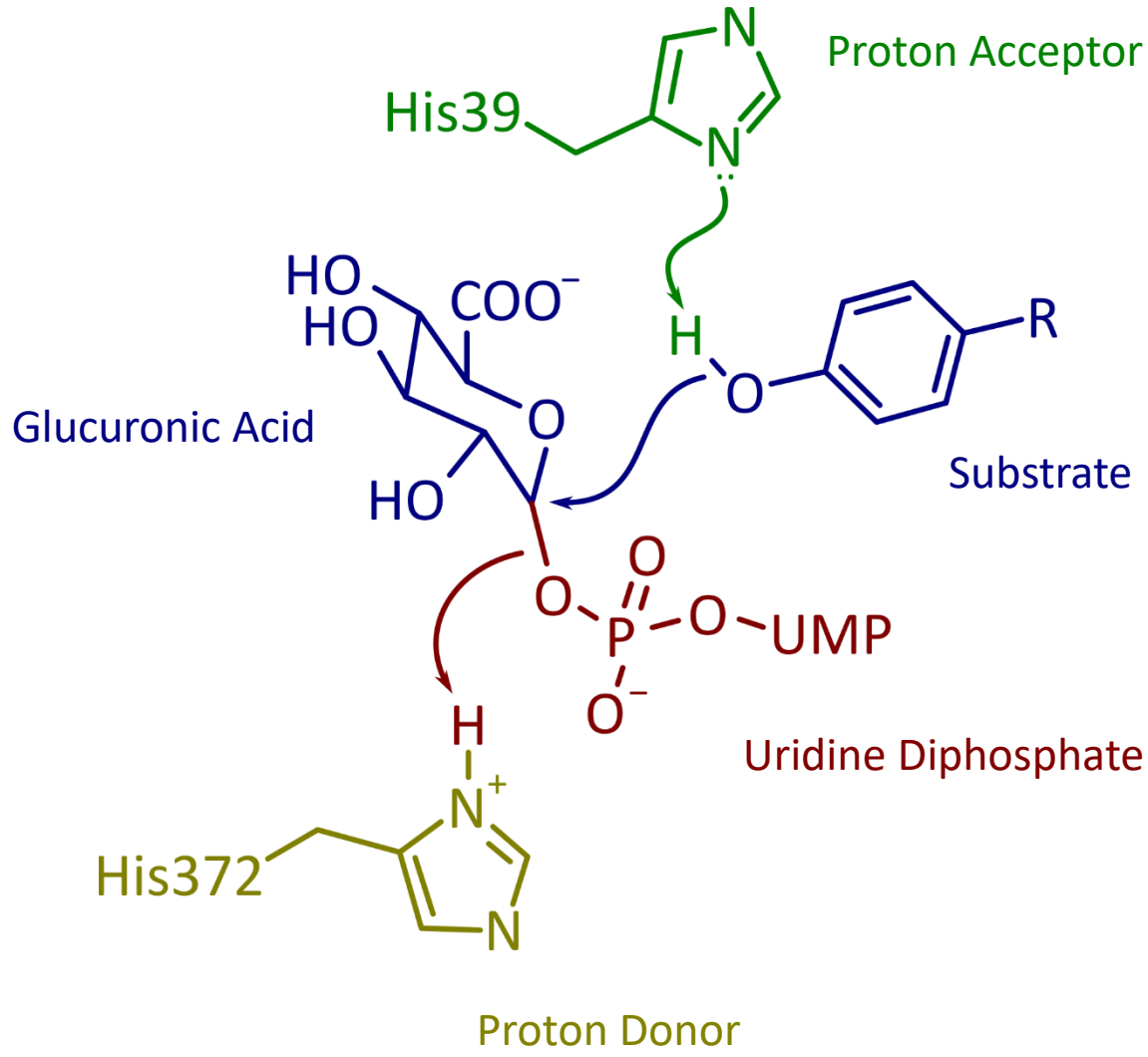
Transition State

- QM simulations to determine reaction mechanism using DFT
- Complex reaction mechanism
  - Proton transfers with active-site histidine residues
- Calculate activation energy, ΔH_A



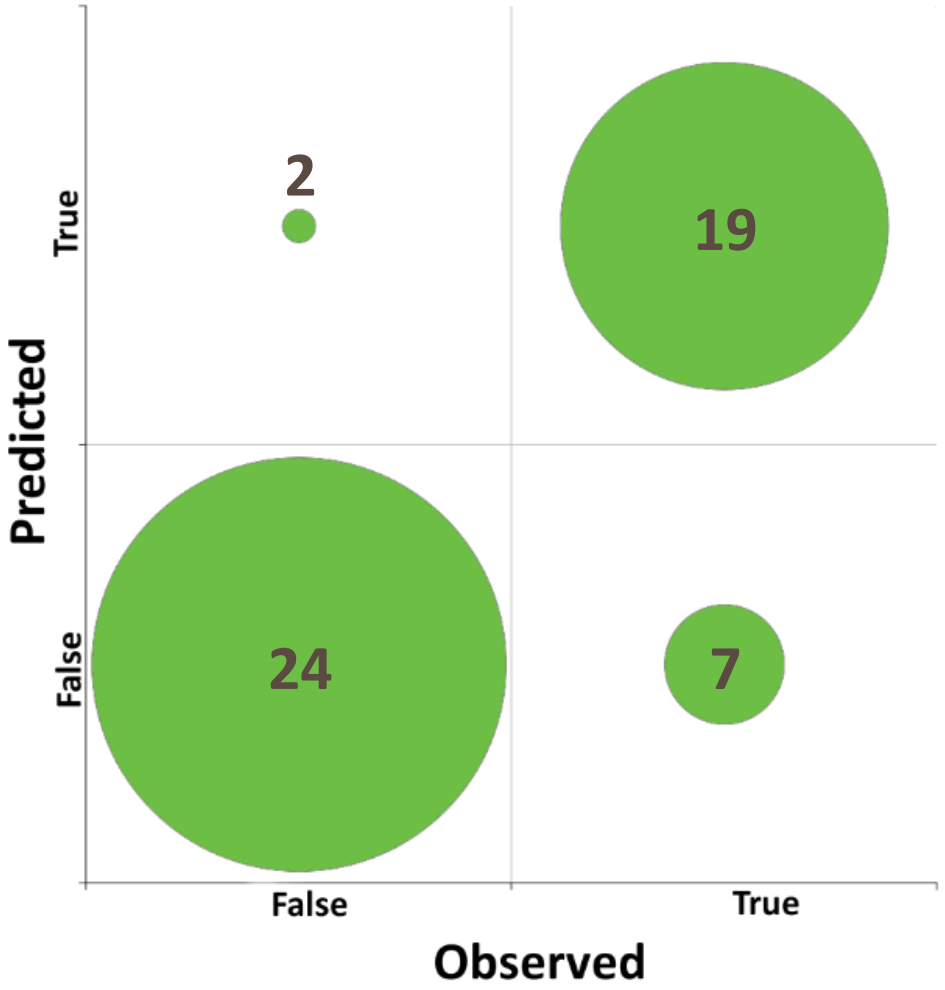
Transition State

- QM simulations to determine reaction mechanism
- Complex reaction mechanism
  - Proton transfers with active-site histidine residues
- Calculate activation energy, ΔH_A



Example – Prediction UGT1A1 Metabolism

- Activation energies calculated with semi-empirical QM model of transition state
- Steric and orientation descriptors included
- Data set
  - 79 molecules
  - 242 potential sites of metabolism
- Gaussian processes machine learning
- Classification of potential sites as metabolised (True) or not (False)
- Results on independent test set
  - Kappa = 0.65
  - Accuracy 83%



Conclusions

- Detailed QM simulations enable us to understand the reaction mechanisms for metabolism
- This enables us to predict metabolism with greater accuracy and transferability
  - Reaction energetics are important factor governing metabolism
  - Combined with steric and orientation effects of protein environment
- Combining models of different steps in the catalytic cycle enable us to predict routes, sites and products of metabolism
  - E.g. WhichP450 and regioselectivity
- For more information
  - J. Tyzack *et al.* (2017) J. Chem. Inf. Model **56**(1) pp. 2180-2193
  - P Hunt *et al.* (2018) J. Comp.-Aided Mol. Des. **32**(4) pp. 537-546
  - matt@optibrium.com

Booth 415

Acknowledgements

- P450 Metabolism
  - John Tyzack
  - Rasmus Leth
  - Many former colleagues from Camitro, ArQule, Inpharmatica and BioFocus
  - This research has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under the grant agreement no 602156

- UGT Metabolism
  - Mario Öeren
  - David Ponting
  - Funding from Lhasa Limited

- FMO Metabolism
  - Peter Walton
  - Mario Öeren

- All of the above – Peter Hunt