On the Cutting-edge of Drug Metabolism

Optibrium Inc.'s President, Dr. Tamsin Mansley, breaks down all you need to know about the challenges of drug metabolism, and the *in silico* approaches to overcome these problems.

Determining metabolic fate is crucial in drug discovery. Horror stories are often shared around expensive, late-stage failures due to unexpected metabolism. Many challenges crop up, including poor metabolic stability, low bioavailability, unforeseen drug-to-drug interactions, issues from genetic polymorphisms, and the formation of reactive or toxic metabolites. Early *in silico* modelling can help to prevent any problems down the line.

Identifying the Enzyme Culprits

The first step to optimise metabolism is understanding which enzymes and isoforms are primarily responsible for your compound's metabolism. Then, you can identify the sites on your molecule that these enzymes are metabolising, and how to design your compound to block these.

There are a range of different enzyme families which may be involved in metabolism. For example, cytochrome P450s, aldehyde oxidases and flavin-containing monooxygenases can cause oxidation of your compound's functional groups. Sulfotransferases and uridine diphosphate glucuronosyltransferases cause conjugation of compounds to polar groups. Additionally, within each enzyme family, there are numerous different isoforms which are functionally similar enzymes that differ slightly in amino acid sequence.

Using classical categorisation models, it is possible to quickly determine which enzyme families and isoforms are most likely to metabolise a specific atomic site.[1–2] This can indicate compounds which can be metabolised by multiple enzymes, with multiple routes of clearance.

There are two main reasons you might want compounds with multiple routes of clearance. Firstly, genetic polymorphisms between individual patients may mean different isoforms of enzymes are present or absent and in varying concentrations. Therefore, in situations where only one isoform is responsible for drug metabolism, issues related to toxic drug build-up may arise in certain populations.

Similarly, single clearance routes increase risks from drug-to-drug interactions. Co-administered drugs may inhibit or induce action by certain drug metabolising enzymes, causing variability in a patient's exposure to the relevant drug. By ensuring multiple routes of drug clearance, these effects can then be mitigated.

Mapping Metabolic Liabilities

Knowing which enzymes cause your compounds' metabolism is only half the battle. To optimise metabolic stability, you also need to identify where your compounds may be metabolised. To model this regioselectivity we can take a dual approach, considering both the reactivity and the accessibility of each atomic site to metabolism.[3–6]

The reactivity of a specific site on a compound to a particular metabolic reaction can be modelled with quantum mechanical simulations. These physics-based methods take a holistic view of a molecule and the electronic distributions within it and hence the electron flow within a reaction pathway. The reactivity of each site on a molecule will be specific to the enzyme family, but will not vary between isoforms of the same enzyme.

The accessibility component of a regioselectivity model is influenced by the substrate's molecular shape and functional groups, along with the particular enzyme's active site structure. This means accessibility will be specific to each isoform and enzyme family. The particular steric and/or polar features within both the enzyme binding pocket and the substrate will determine the substrate's orientation and whether a particular area can access the active site; thus, some sites will be less vulnerable to metabolism than others. Accessibility effects can be modelled using descriptors rooted on each site of metabolism on the ligand.

Reactivity and accessibility effects for each enzyme and isoform can be combined using robust machine learning models, trained on high-quality data and tested on an independent test set. By applying these regioselectivity models, a good comprehension can be gained into which labile sites need to be blocked for increased metabolic stability.

Too Many Metabolites

In addition to knowing if your structure will be metabolised, you need to be able to understand which compounds will form. Will any toxic or reactive metabolites be generated, causing any serious adverse side effects? Unfortunately, metabolite prediction is not as easy as it may seem.

Any predictive models must satisfy two criteria. Firstly, are the structures generated accurate representations of potential metabolites? Are all the experimentally-observed compounds being predicted? This can involve an added layer of complexity when metabolites are unstable and quickly undergo further reactions such as hydrolysis.

Secondly, and much more challenging, is considering the number of potential metabolites being predicted. Of course, you can generate every possible structure which a compound may be metabolised into, but how easy is it to pick out the subset of metabolites that are actually observed experimentally? Does your method suffer from vast over-prediction, obscuring the important, relevant experimentally observed metabolites?

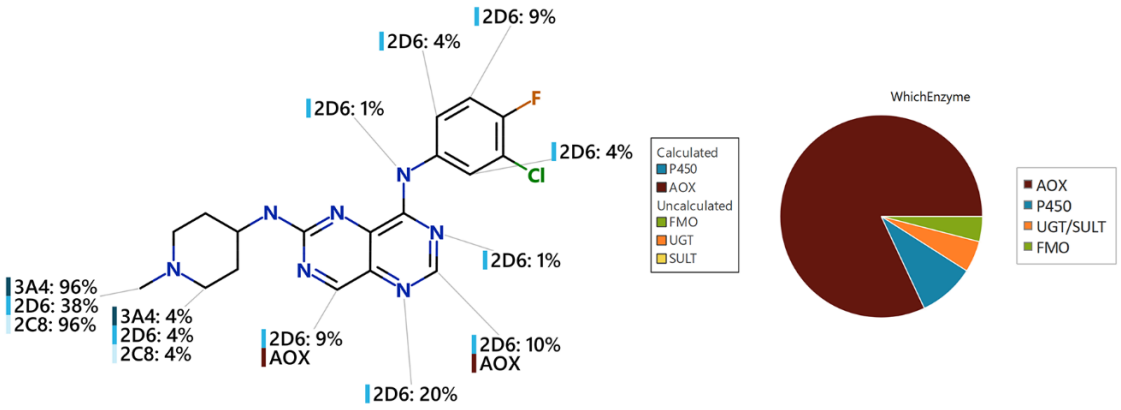

Figure 1: The output of the StarDrop™ Metabolism module for Falnidamol. Here the WhichEnzyme™ pie chart clearly shows AOX as the major metabolising enzyme.

Using a heuristics approach to combine regioselectivity models and classification models, it is possible to predict the most likely metabolites with a much higher sensitivity than traditional rule-based methods. This streamlines the discovery process, minimising time wasted sifting through irrelevant metabolite predictions and making interpretation of metabolite ID experiments much easier.

Picking the Best Animals for the Job

One final common query around drug metabolism is relating to preclinical studies. Selecting inappropriate animal species for these studies can have devastating consequences. Take, for example, the case of Falnidamol. This was a clinical drug candidate under investigation as a cancer treatment, which passed routine pharmacokinetics studies in rats and dogs, but failed in clinical trials, due to extremely low oral exposure in humans. This was due to rapid metabolism by AOX, which was not picked up during the preclinical PK studies, as rats have low aldehyde oxidase activity and dogs are devoid of aldehyde oxidase. More appropriate preclinical species for accurate preclinical trials could have been guinea pig or rhesus monkey, which have high aldehyde oxidase activity.

Had appropriate early-stage modelling been used, the right species could have been selected and labile sites identified and blocked to improve metabolic stability. This would have saved enormous amounts of time, money, and resources for the researchers.

Modelling animal species is more difficult than modelling human enzyme regioselectivity, due to the limited quality data available with which to build models. However, there are models currently available for rat, mouse and dog cytochrome P450s, with the potential for future research pathways into different animal species or enzyme models.

A Future Perspective

There are a few limitations around what we can currently achieve with this type of modelling. The first is data. High quality enzyme specific data is needed to train relevant models, so not every enzyme family or isoform, and not every common preclinical species can be modelled right now. As research continues and data improves, coverage of models can become more comprehensive.

Modelling metabolism using quantum mechanical simulations is also very computationally expensive and time intensive. New methodology is constantly being developed to improve this, for example machine learning interatomic potentials.[7] This will speed up future calculations to make metabolism prediction easier and more commonplace in discovery.

The extent of our current abilities in metabolism prediction has grown exponentially in recent years, with medicinal chemistry teams and DMPK scientists well-supported to create safe, efficacious drugs. As this field continues to progress, the future is looking even brighter.

REFERENCES

1. M. Öeren *et al.*, "Predicting routes of phase I and II metabolism based on quantum mechanics and machine learning", Xenobiotica, 2023, 1–49
2. P. A. Hunt *et al.*, "WhichP450: a multi-class categorical model to predict the major metabolising CYP450 isoform for a compound", J Comput.-Aided Mol. Des., 2018, 32, 537–546
3. M. Öeren *et al.*, "Predicting regioselectivity of AO, CYP, FMO, and UGT metabolism using quantum mechanical simulations and machine learning", J. Med. Chem. 2022, 65, 20, 1406–1408.
4. M. Öeren *et al.*, J. Comput.-Aided Mol. Des., 2021, 35, 4, 541-555
5. M. Öeren *et al.*, "Predicting regioselectivity of cytosolic sulfotransferase metabolism for drugs", J. Chem. Inf. Model. 2023, 63, 11, 3340–3349
6. J. D. Tyzack *et al.*, "Predicting reactivity to drug metabolism beyond P450s - modelling FMOs and UGTs", J. Chem. Inf. Model. 2016, 56, 11, 2180–2193
7. E. Gelžinytė *et al.*, "Transferable Machine Learning Interatomic Potential for Bond Dissociation Energy Prediction of Drug-like Molecules", J. Chem. Theory Comput. 2024, 20, 1, 164–177

Dr. Tamsin Mansley

Dr. Tamsin Mansley, President, Optibrium Inc. and Head of Application Science, Optibrium. Tamsin holds a PhD in Organic Chemistry from the University of East Anglia in the UK and pursued postdoctoral studies in the labs of Prof. Philip Magnus at the University of Texas, Austin. She is an experienced drug discovery scientist with experience as a medicinal chemist in pharma and biotech and has worked in computational chemistry, supporting discovery organisations for over 20 years. Her interests lie in coupling machine learning and artificial intelligence techniques with generative chemistry approaches to guide decisions and accelerate discovery projects.