Review

# Computational tools for modeling xenometabolism of the human gut microbiota

**Martina Klünemann, Melanie Schmid, and Kiran Raosaheb Patil**

Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

**The gut microbiota is increasingly being recognized as a key site of metabolism for drugs and other xenobiotic compounds that are relevant to human health. The molecular complexity of the gut microbiota revealed by recent metagenomics studies has highlighted the need as well as the challenges for system-level modeling of xenobiotic metabolism in the gut. Here, we outline the possible pathways through which the gut microbiota can modify xenobiotics and review the available computational tools towards modeling complex xenometabolic processes. We put these diverse computational tools and relevant experimental findings into a unified perspective towards building holistic models of xenobiotic metabolism in the gut.**

## Gut microbiota as a site of xenometabolism

The gut microbiota has been shown to modify or metabolize several kinds of xenobiotics, from novel cancer drugs through millennia old analgesics to dietary components [1–7]. Recent studies have also highlighted the feasibility of exploiting and manipulating this microbe-mediated xenometabolism to improve the host health or to prohibit medicinal side effects. For example, Wallace *et al*. [6] recently showed that a deleterious biotransformation of the cancer drug irinotecan can be averted by inhibiting bacterial β-glucuronidase. On a more general level, probiotic bacteria like *Lactobacillus* sp. have been shown to ease *Clostridium difficile*-associated diseases, diarrhea, and other side effects of antibiotics [8,9].

Owing to the advances in various omics technologies, molecular pathways of xenometabolism in the gut microbiota have now started to unfold through the identification of responsible microorganisms and enzymes [6,10,11]. With the help of metagenomics tools, it is now possible to determine the identity of a large fraction of the microbial species colonizing the human gut [12,13]. These tools are also revealing the genetic repertoire of the gut microbiome

*Corresponding author:* Patil, K.R. (patil@embl.de).
*Keywords:* metabolic modeling; microbial communities; functional metagenomics; biotransformation.

0167-7799/$ – see front matter
 http://dx.doi.org/10.1016/j.tibtech.2014.01.005

## Glossary

**BRENDA:** Braunschweig Enzyme Database – one of the main collections of enzyme function and activity data. The database contains several xenobiotic–enzyme interactions (for examples, see Figure 2).
**Constraint-based modeling:** a metabolic modeling technique that uses mass balance, reaction directionality and metabolite uptake/secretion constraints to estimate intracellular fluxes (reaction rates) for a given metabolic network.
**Enterohepatic cycle:** circulation of native or xenobiotic compounds between liver and gastrointestinal tract. Hepatic biotransformations that increase water solubility of the compounds, for example, glucuronidation of bilirubin, is a common denominator for most known enterohepatic circulations.
**Enzyme promiscuity:** property of enzymes whereby reactions are catalyzed with a degree of non-specificity. A promiscuous enzyme can act on multiple substrates and/or produce different products starting with the same substrate.
**Genome-scale metabolic model:** a comprehensive mathematical representation of metabolic capabilities of a cell. A typical bacterial model consists of a network of several hundreds of reactions that are compiled based on genomic, biochemical and other evidences.
**KEGG:** Kyoto Encyclopedia of Genes and Genomes. A comprehensive database of high-level gene functions and their organization in a pathway or cellular context. The well-maintained and annotated pathway maps, reaction modules, and drug database of KEGG are of particular interest for modeling enzyme promiscuity and predicting xenometabolic pathways.
**Machine learning algorithms:** computational methods aimed at prediction based on patterns learned from a given input dataset (training data). In the absence of mechanistic models, machine learning is a useful tool for predicting whether a xenobiotic compound of interest would be susceptible for chemical modification by a particular enzyme. Known enzyme–compound relations from databases like BRENDA, KEGG, and UM-BBD can be used as training data.
**Phase I and II metabolism:** metabolic reactions that predominantly occur in the liver and gastrointestinal epithelial tissue. Whereas the phase I enzymes introduce reactive or polar chemical groups in the xenobiotic compounds, the phase II enzymes catalyze conjugation reactions. Together, phase I and II metabolism alters the activity of xenobiotics and/or converts them into more water-soluble compounds, often leading to detoxification and eventual excretion.
**Structure-activity relation (SAR) paradigm:** SAR paradigm assumes that similar molecules have similar biological activities. Thus, the activity of a xenobiotic compound of interest can be described based on the activity of known compounds with similar physicochemical/molecular properties. Different biochemical activities depend on different physicochemical properties, therefore, the most critical step in the application of SAR paradigm is to identify a similarity metric that is most appropriate for the activity in question. A quantitative approach to structure–activity relation (QSAR) allows predicting the degree of expected activity based on the degree of similarity.
**Site of metabolism:** site in a chemical compound at which the major chemical change takes place during a given biochemical transformation.
**Thermodynamic feasibility analysis:** feasibility analysis for a particular biochemical transformation based on estimation of the overall change in the Gibb's free energy (ΔG) accompanying the reaction. Only reactions with negative ΔG values are thermodynamically feasible.
**UM-BBD:** University of Minnesota Biocatalysis/Biodegradation Database; contains extensive data on microbial biocatalytic reactions and biodegradation pathways.
**Xenobiotics:** molecules of foreign origin encountered by the body, such as drugs and dietary compounds that are not naturally produced by the human body.
**Xenometabolism:** enzyme-mediated biochemical transformation/degradation of xenobiotics.

in unprecedented detail. The resulting rich datasets are enabling the characterization of the gut microbial communities and their association with health [14]. Metagenomics datasets are also providing a starting point for modeling the collective metabolic behavior of the gut microbiota [15,16]. In parallel to these advances stemming from metagenomics studies, more experimental evidence is piling up supporting the key role of gut microbiota in xenometabolism [3–5]. Many of the aforementioned studies have gained useful insight from or were validated in animal models. Indeed, animal models have been fundamental for understanding the biology of gut microbiota in general [17]. Complementing metagenomics, metabolomics has made it possible to trace the metabolic fate of xenobiotic compounds [18–20], leading the recent resurge in the research on xenometabolism [21–23].

## Role of modeling in tackling complexity of xenometabolic processes

Xenometabolic processes in the gut can be highly complex due to three main reasons: the widespread promiscuity of metabolic enzymes; the compositional complexity of the gut microbiota; and the interactions between the host and the microbe-mediated xenometabolism. The promiscuity of metabolic enzymes [24–26] means that the number of possible routes through which a xenobiotic compound can be metabolized or modified increases combinatorially with the enzymatic repertoire of the microbiota. The compositional diversity and spatial heterogeneity of the microbiota and the host–microbiota interaction through the enterohepatic cycle add another layer to this complexity. These inherent complexities and the limitations in obtaining *in vivo* data from human subjects raise the appeal for computational modeling of xenometabolism. A holistic modeling platform accounting for combined host–microbiota xenometabolism can be instrumental in drug development and for devising personalized medicine strategies. Such a platform would allow prediction of potential xenometabolic pathways and thereby generation of testable hypotheses. Although no single platform currently tackles all of the distinct challenges in modeling xenometabolism, several tools are available that are capable of addressing the key individual steps. These tools range from prediction of enzyme-level biotransformation [27–30] to

**Table 1. Representative computational tools relevant for modeling xenometabolism by the gut microbiota**

| Tool | Original scope | Main features/techniques | Web access/ open source | Refs |
|---|---|---|---|---|
| **Single-step biotransformation prediction** | | | | |
| MetaPrint2D | Sites of human phase I metabolism | Comparison to reaction centers from the proprietary Symyx Metabolite database | Yes/yes[a] | [30] |
| MetaPrint2D-React | Subset of biotransformations in phase I metabolism | Based on Metaprint2D | Yes/– | [30] |
| Meteor | Metabolites of phase I and phase II metabolism | Reaction rules based on multiple knowledge bases | No/no | [44] |
| MaRIboES | Generic | Generalized reactions based on BRENDA | No/yes | [28] |
| Chen, Feng *et al.* | Generic | Machine learning | No/upon request | [48] |
| Mu, Unkefer *et al.* | Sites of enzymatic attack | Reaction rules based on KEGG; machine learning | Yes/no | [50] |
| **Biotransformation pathway prediction** | | | | |
| MetabolExpert | Biodegradation by plants and animals | Knowledge based reaction rules | No/no | [45] |
| UM-PPS | Biodegradation by aerobic bacteria | Reaction rules based on UM-BBD; expert knowledge | Yes/no | [29] |
| CRAFT | Biodegradation by aerobic bacteria | Reaction rules based on UM-BBD | No/yes | [b] |
| BNICE | Generic | Generalized reactions based on EC numbers; thermodynamic feasibility analysis | No/no | [27] |
| PathPred | Microbial biodegradation | Machine learning based on KEGG | Yes/no | [47] |
| Desharky | Microbial biodegradation to host native metabolites | Pathway search using KEGG; host-organism specific | No/yes | [85] |
| **Species metabolic modeling** | | | | |
| Model SEED | Bacterial genome-scale metabolic models | Constraint-based modeling; automated model construction | Yes/yes | [57] |
| COBRA Toolbox | Generic | Constraint-based modeling; MATLAB and Python based | No/yes[c] | [61] |
| OptFlux | Generic | Constraint-based modeling; Java based | No/yes | [86] |
| **Microbial community modeling** | | | | |
| OptCom | Generic; continuous growth | Constraint-based modeling; accounts for growth requirements of individual species | No/academic use | [31] |
| Klitgord and Segrè | Induction of cooperation in communities | Constraint-based modeling; focus on nutritional composition of growth media | No/yes | [32] |
| Freilich *et al.* | Competitive and cooperative interactions | Constraint-based modeling; pairwise interactions | No/– | [84] |

Abbreviations: EC, Enzyme Commission; –, no information found.

[a]Tool is open source, database is excluded.

[b]http://www.molecular-networks.com/products/craft.

[c]Works within the licensed MATLAB framework.