Genome-scale metabolic reconstruction of *Escherichia coli* to account for promiscuous enzyme activity and underground metabolism

Maurício Alexander de Moura Ferreira[1], Eduardo Luís Menezes de Almeida[1], Wendel Batista da Silveira[1], Zoran Nikoloski[2,3,*]

[1] *Department of Microbiology, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil.*
[2] *Bioinformatics, Institute of Biochemistry and Biology, University of Potsdam, Potsdam, Germany.*
[3] *Systems Biology and Mathematical Modelling, Max Planck Institute of Molecular Plant Physiology, Potsdam, Germany.*
** Corresponding author. E-mail: zoran.nikoloski@uni-potsdam.de*

Many enzymes display side activities by catalysing reactions other than their main reaction. These enzymes are known as promiscuous enzymes and form an alternative metabolic network of reactions termed underground metabolism. While many of these reactions are physiologically irrelevant given the low enzymatic activity, they can act as a reservoir of novel enzyme functions that can be exploited for biotechnological and evolutionary purposes. Protein-constrained genome-scale metabolic models (pcGEMs) provide the means to integrate promiscuous functions and enzyme usage allocated to them. However, it is challenging to integrate kcat values for promiscuous enzymes in pcGEMs, since (i) promiscuous enzymes affect the gene-protein-reaction (GPR) rules and (ii) databases may contain information about non-canonical enzyme/substrate pairs. Here, we perform an investigation on how kcat values of underground metabolism can be obtained and integrated into pcGEMs. To this end, we manually curated the *Escherichia coli* GEM iML1515 and included predicted underground reactions. The curated model, which we term iML1515u, contains 3046 reactions, of which 695 are underground, including transport and exchange reactions as well as substrate-product conversions. These reactions also account for the production of 160 different value-added chemicals, such as polylactic acid, 3-hydroxypropanoate and 3-hydroxybutryolactone, used for bioplastics production; and hydroquinone, (R)-3-hydroxybutanoate and 4-hydroxy-3-methoxy-benzaldehyde, used for the pharmaceutical industry. Next, we retrieved the kcat values for all enzymes in the model from the BRENDA database and compared them to those predicted by a data-driven approach termed DLKcat. We found that the kcat values retrieved from BRENDA are generally larger than those predicted by DLKcat, with implications to enzyme allocation to satisfy cellular demands. Given that the BRENDA database does not account for all possible promiscuous enzyme activities, many of the kcat values were obtained by similarity to other enzymes from the same catalytic group, usually obtained for other substrate-enzyme pairs; in contrast, the kcat values predicted by DLKcat are obtained by pairing an enzyme sequence and a substrate, generating a kcat value that is tailored to the enzyme-substrate pair. Lastly, we modified existing approaches to predict enzyme allocation, making sure that a subpool allocation to a promiscuous enzyme is used for a particular reaction. The generated resources will facilitate investigation of underground metabolism in the context of pcGEMs.

Key words: Enzyme promiscuity; Metabolic modelling; Catalytic rates; Enzyme allocation.

Reconstrução da rede metabólica de *Escherichia coli* incorporando atividade enzimática promíscua e o metabolismo underground

Muitas enzimas exibem atividades secundárias, catalisando reações diferentes de sua reação principal. Essas enzimas são conhecidas como enzimas promíscuas e formam uma rede metabólica alternativa de reações chamada metabolismo *underground*, servindo como um reservatório de novas funções enzimáticas que podem ser usadas para fins biotecnológicos e evolutivos. Este estudo investiga como os valores de eficiência catalítica podem ser obtidos e integrados em modelos metabólicos em escala genômica com restrições enzimáticas para estudo do metabolismo *underground*. Além disso, propomos um método baseado em otimização matemática para predizer fluxos metabólicos nestas reações, dividindo o uso de enzimas em subconjuntos.

Palavras-chave: Promiscuidade enzimática; Modelagem metabólica; Eficiência catalítica; Alocação de enzimas.

Acknowledgements: *This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.*