Representation of ME-models in SBML

Introduction

Models of metabolism and macromolecular expression (ME-models)[1] are a recent extension of the well-established genome-scale models of metabolism (M-models). They provide increased predictive capabilities and accuracy when compared to M-models at the expense of a largely increased size and complexity. Until now, ME-models cannot be encoded in the SBML, which is the *de-facto* standard format for storing and communicating models of biological systems between different tools and software environments.

Aims

The aim of this master thesis is to develop and suggest an extension of SBML to encode ME-models. The planned work comprises the following specific aims:

- Study the structure and requirements of ME-models
- Compare these structures with the current version of SBML including its packages
- Identification of requirements and additional data structures in SBML or attributes for existing SBML constructs needed for encoding ME-models
- Discussion of identified requirements and proposed SBML data structure with ME-modelers in order to create a consensus version
- Prototyping an extended SBML version in JSBML (the Java™ API for working with SBML)
- Implementation of a bidirectional converter between the JSON format that has been developed for the COBRAme software[2] and SBML in Java™ using the Jackson library with focus on performance because of the enormous size of ME-models (fig. 1)
- Communication of findings to the COBRA and SBML communities in order to either propose an entirely new package for SBML or an updated version of the FBC package (for flux balance constraints)

In order to create an SBML extension tailored to the needs of ME-modelers this project will be conducted in close cooperation with researchers from the University of California, San Diego, who created the COBRAme software[2]. In the course of the project it would be ideal to reach out to further groups who actively develop ME models in order to ensure that the proposed SBML extension will cover different branches of ME-models.

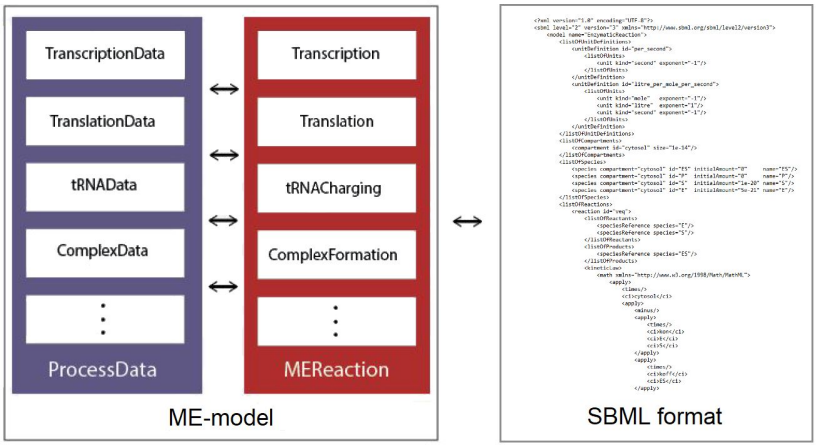

Figure 1: Bidirectional conversion between the ME-model representation in COBRAme and the representation in SBML.

References

1. Lerman, J. A. *et al.* In silico method for modelling metabolism and gene product expression at genome scale. *Nat. Commun.* **3,** 929 (2012).
2. Lloyd, C. J. *et al.* COBRAme: A Computational Framework for Building and Manipulating Models of Metabolism and Gene Expression. *bioRxiv* 106559 (2017). doi:10.1101/106559