# Biochemical Space: A Framework for Formal Description and Annotation of Complex Biological Processes

Matej Troják, David Šafránek, Jakub Šalagovič,
Františka Romanovská, and Matej Hajnal

*Faculty of Informatics, Masaryk University*
*Brno, Czech Republic*

**Abstract**

Biochemical Space is a novel framework for practical description of complex biological processes. It is developed as a part of Comprehensive Modeling Platform – a web-based platform for modelling and analysis of biological processes. In the context of the platform, Biochemical Space binds complicated quantitative models with an easy-to-understand, yet formal and compact qualitative description. It allows us to specify formal and well-annotated reaction networks of chemical entities and elemental reactions onto which the mathematical models are projected. Biochemical Space is supported by Biochemical Space Language that combines state-of-the-art rule-based techniques with metadata formats developed in well-known annotation databases.

## 1 Background

The concept of Biochemical Space (BCS) forms a crucial part of *Comprehensive Modeling Platform* (CMP) – a general platform for computational modelling and analysis of biological processes (see Fig. 1). CMP was first introduced in [7] as a concept for formal representation of internally consistent reduced models of oxygenic photosynthesis [8] and further refined to a general platform in [5].

BCS avoids quantitative issues and focuses solely on compact representation of biological reactions and their annotation [6] (see Fig. 3). This representation is treated by adopting selected concepts from state-of-the-art rule-based languages (BNGL [3], Kappa [1]). BCS also reflects SBML level 3 [4] in generalising of compartments in terms of a hierarchy of locations, introducing entity states, and dealing with related combinatorial explosion.

---

[1] This work has been supported by the Czech Research Infrastructure for Systems Biology grant No. LM2015055.

The CMP supported with BCS provides a unique solution based on integrating the well-acknowledged systems biology standards with advanced computational techniques such as static and dynamic analysis.

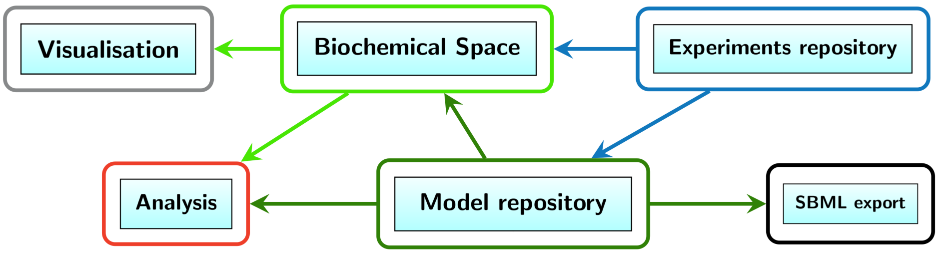

Fig. 1. Biochemical Space in the context of the modelling platform CMP. Modules of experiments and models are connected via common Biochemical Space. It allows us to apply indirect analysis and visualisation of the modules. Moreover, SBML export of the models with mapping on BCS is supported.

## 2 Framework Description

Biochemical Space is a general annotation format for description of reaction networks including textual annotation and links to existing annotation databases. The rigorous (rule-based) core of BCS is called *Biochemical Space Language* (BCSL) and is made by declaration of chemical entities and reaction rules. Formal definition of the language including operational semantics is described in [2]. BCSL has the following features:

- description of rules/reactions including stoichiometry;
- definition of states encoding different forms of an entity;
- definition of composite entities in an abstract way by means of coexistence of individual subentities rather than by enumerating precisely the individual bonds;
- partial definition of internal entity structure without the need to enumerate all possibilities (e.g., a serine-phosphorylated protein);
- spatial organisation of entities in terms of hierarchical locations;
- variables minimising the need for repetitive definitions.

Employing BCSL, biological processes are formalised and mathematical models are mapped and annotated (see Fig. 2).

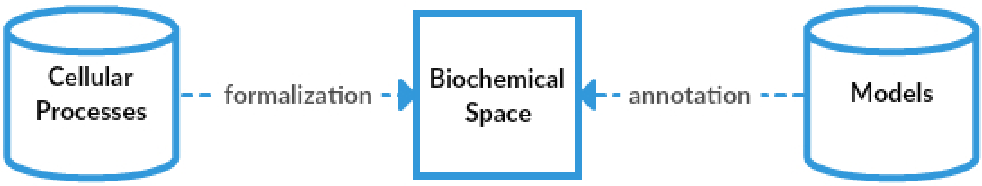

Fig. 2. The general role of BCS is to decrease the space between biological and mathematical domains. It is done by formalization of biological description while annotating models.

An example of existing application of Biochemical Space for modelling of cyanobacteria processes is available at `e-cyanobacterium.org`. In the current version (July 2017), the following processes of cyanobacteria are covered: environmental processes, respiration and photosynthesis, carbon concentrating mechanism, circadian clock, and metabolism.

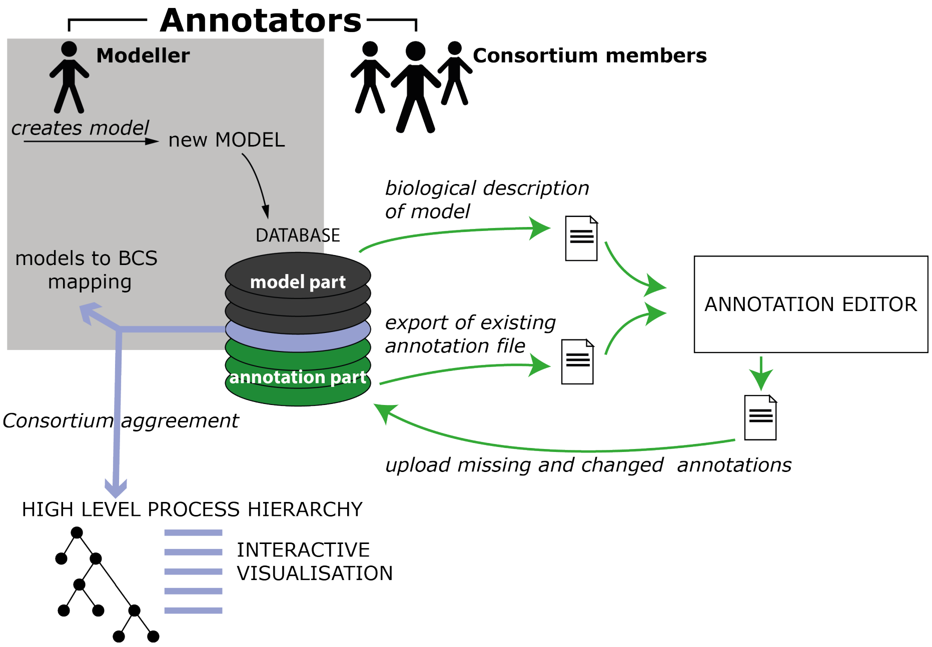

Fig. 3. Annotation is one of the key attributes of BCS. The process of annotating in not trivial and requires precise attention and cooperation of people from multiple fields of study.

## 3 Conclusions

In conclusion, the main goal of BCS as a part of CMP is to simplify model building tasks by providing simple and clear way of notation easily understandable by both modellers and biologists.

## References

[1] Danos V. et al., *Formal molecular biology*, Theor. Comput. Sci. **325** (2004), pp. 69–110.

[2] Děd T. et al., *Formal Biochemical Space with Semantics in Kappa and BNGL*, Electronic Notes in Theoretical Computer Science **326** (2016), pp. 27 – 49.

[3] Faeder J. R. et al., *Rule-Based Modeling of Biochemical Systems with BioNetGen - Springer*, number 500 in Methods in Molecular Biology, Humana Press, 2009 .

[4] Hucka M. et al., *The Systems Biology Markup Language (SBML): Language Specification for Level 3 Version 1 Core*, Journal of integrative bioinformatics (2010).

[5] Klement M. et al., *A comprehensive web-based platform for domain-specific biological models*, in: *Proceedings of the fourth International Workshop on Interactions between Computer Science and Biology (CS2Bio'13)* (2013), pp. 61–67.

[6] Le Novere N. et al., *Minimum information requested in the annotation of biochemical models (MIRIAM)*, Nat Biotech **23** (2005), pp. 1509–1515.

[7] Nedbal L. et al., *Scaling and integration of kinetic models of photosynthesis: Towards comprehensive e-photosynthesis*, in: *Photosynthesis in silico*, Advances in Photosynthesis and Respiration **29**, Springer, 2009 pp. 17–29.

[8] Šafránek D. et al., *E-photosynthesis: Web-based platform for modeling of complex photosynthetic processes*, Biosystems **103** (2011), pp. 115–124.