Executable Biochemical Space for Specification and Analysis of Biochemical Systems*

Matej Troják[1], Jakub Šalagovič[2], David Šafránek[1], Jan Červený[2], Luboš Brim[1], and Matej Hajnal[1]

[1] Systems Biology Laboratory, Masaryk University, Brno, Czech Republic
[2] Global Change Research Centre AS CR, v. v. i., Brno, Czech Republic

Abstract. Biochemical Space Language is a rule-based high-level language for hierarchical description of biological structures and mechanistic description of chemical reactions by means of compact rules. Its uniqueness is in the possibility of creating hierarchically related objects and describing their interactions them on different levels of abstraction while preserving human-readable notation. Moreover, it makes an inseparable part of general semi-formal description format able to capture annotation information along with structural details.

Keywords: rule-based language, formal specification, annotation

1 Introduction

We present a work in progress on a rule-based language called *Biochemical Space Language* (BCSL) [3]. It forms an integral part of a general framework called *Comprehensive Modelling Platform* [6], developed for comprehensive modelling in systems biology.

The general goal of the platform is to respect the need for maintaining existing ODE models but allows to align them with a mechanistic rule-based description that is understandable by biologists, compact in size, executable in terms of allowing basic analysis tasks ensuring consistency of the description, and provides links to existing bioinformatics annotation databases. Such a comprehensive solution allows to support effort of modellers in building mathematical models that have clear biochemical meaning and can be easily integrated. Moreover, mechanistic descriptions can be later used as computational models having all advantages of rule-based modelling.

Although rule-based models make a great alternative to mathematical models, they are not yet sufficiently used in practice. The reason is that existing formalisms rely on cryptic (symbolic) syntax and they are limited to a specific subset of interactions or are too abstract [2,4]. These formalisms can be thus understood as low-level formats. Several high-level frameworks have been developed based on principles of these formalisms: rxncon [9] focuses on regulatory interactions and allows construction of rules from experimental evidence,

* This work has been partially supported by the Czech Science Foundation grant No. 18-00178S

LBS [8] enrich rule-based framework with modularity, PySB [7] embeds BNGL into Python, BioCHAM [1] explicitly separates rules from their mathematical semantics. None of these frameworks provides a sufficiently universal solution for description and annotation of heterogeneous biophysical processes integrated at the cellular level.

Apparently, different approaches need to be combined accordingly to make a universal hierarchical modelling and annotation base that supports executability. To that end, we developed a semi-formal description format called *Biochemical Space* (BCS) [5], containing both annotation (external links and informal description) and descriptive part in terms of BCSL.

2 Language description

BCSL is a high-level language for hierarchical description of biological structures and mechanistic description of chemical reactions by means of compact rules. BCSL relies on the formal basis of the rule-based methodology while preserving user-friendly syntax (for example of a rule, see Figure 1). The general goal of the language is to deal with combinatorial explosion of numerous interactions by providing a concise and understandable notation. The language combines the following aspects:

- *human-readability* – easy to read, write, and maintain;
- *executability* – formal executable semantics is defined allowing efficient static analysis and consistency checking;
- *universality* – principally different cellular processes can be sufficiently combined in a single specification;
- *scalability* – combinatorial explosion of the description is avoided;
- *hierarchy* – object types are described hierarchically allowing compositional assembly from simpler structures.

Rule ID:	*KaiC6* serine (de)phosph.
Rule equation:	$Ser\{u\}::KaiC::KaiC6::cyt \Leftrightarrow Ser\{p\}::KaiC::KaiC6::cyt$
Rule name:	*KaiC6* serine phosphorylation and dephosphorylation
Classification:	phosphorylation, dephosphorylation
Description:	*KaiC* molecule is phosphorylated/dephosphorylated on serine amino acid residue in *KaiC* hexamer.
Links:	doi::10.1016/j.febslet.2009.11.021

Fig. 1. Example of a complete information about a rule. The record contains annotation information such as *Links* to external sources, *Classification* of described process, and *Description* of the process. *Rule equation* contains a rule written in BCSL employing object state change. Particularly, the rule describes change of state in protein complex of six *KaiC* proteins.

Since the language is rule-based, it has ability to capture described processes in a very concise way. However, the high level of abstraction can cause confusions

in understanding of complicated expressions. In order to improve the clarity of the format and its usability, the language has the following features: description of rules including stoichiometry, which shortens the notation of repetitive objects; definition of states encoding different forms of a object, which decreases the number of distinct objects; definition of composite objects in an abstract way by means of coexistence rather than by enumerating precisely the individual bonds, which decreases combinatorial explosion in terms of the number of possible spatial conformations; partial definition of internal structure without the need to enumerate all possibilities, which is the typical feature for huge biochemical objects changing only in a fraction of its structure; variables minimising the need for repetitive definitions.

BCSL is currently implemented in e-cyanobacterium.org [10], a web-based platform for the modelling and analysis of biological processes occurring in cyanobacteria. The platform provides several features that contribute to the production and presentation of models targeting cyanobacteria. The principal effort is to interlink biological knowledge of cyanobacteria with the benefits of computational systems biology tools.

References

1. Calzone, L., Fages, F., Soliman, S.: BIOCHAM: An Environment for Modelling Biological Systems and Formalizing Experimental Knowledge. Bioinformatics **22**, 1805–1807 (2006)
2. Danos, V., Laneve, C.: Formal molecular biology. Theoretical Computer Science **325**, 69–110 (2004)
3. Děd, T., Šafránek, D., Troják, M., Klement, M., Šalagovič, J., Brim, L.: Formal Biochemical Space with Semantics in Kappa and BNGL. Electronic Notes in Theoretical Computer Science **326**, 27–49 (2016), the 6$^{\text{th}}$ International Workshop on Static Analysis and Systems Biology, SASB 2015
4. Faeder, J.R., Blinov, M.L., Hlavacek, W.S., et al.: Rule-based Modeling of Biochemical Systems With BioNetGen. Methods Mol Biol **500**, 113–167 (2009)
5. Klement, M., Děd, T., Šafránek, D., Červený, J., Müller, S., Steuer, R.: Biochemical Space: A Framework for Systemic Annotation of Biological Models. In: CS2Bio 14. pp. 31–44. Elsevier, Amsterdam (2014)
6. Klement, M., Šafránek, D., Děd, T., Pejznoch, A., Nedbal, L., Steuer, R., Červený, J., Müller, S.: A Comprehensive Web-based Platform for Domain-specific Biological Models. In: Proceedings of the 4$^{\text{th}}$ International Workshop on Interactions between Computer Science and Biology. pp. 61–67. Elsevier (2013)
7. Lopez, C.F., Muhlich, J.L., Bachman, J.A., Sorger, P.K.: Programming Biological Models in Python Using PySB. Molecular systems biology **9**(1) (2013)
8. Pedersen, M., Plotkin, G.: A Language for Biochemical Systems: Design and Formal Specification. In: Transactions on Computational Systems Biology XII: Special Issue on Modeling Methodologies, pp. 77–145. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
9. Romers, J.C., Krantz, M.: rxncon 2.0: A Language for Executable Molecular Systems Biology. bioRxiv (2017)
10. Troják, M., Šafránek, D., Hrabec, J., Šalagovič, J., Romanovská, F., Červený, J.: E-cyanobacterium.org: A Web-based Platform for Systems Biology of Cyanobacteria. In: Proceedings of 14$^{\text{th}}$ International Conference on Computational Methods in Systems Biology (CMSB). pp. 316–322. Springer International Publishing (2016)