Scalable Methods for Modeling Complex Cell Signaling Systems

Ty M. Thomson[1,2], Joshua Havumaki[1], Isha Antani[1], and Gordon Webster[1]

***Short Abstract* — The size and complexity of many complex signaling systems can overwhelm commonly used reaction-based modeling techniques, making it impossible to build and simulate detailed models of these systems without making unwarranted assumptions. To facilitate modeling of complex signaling systems, we developed a novel rule-based modeling framework built on a formal semantic language, Kappa. The stochastic and deterministic simulators that we created avoid enumeration of all possible complexes, enabling simulation of models that could not otherwise be studied. These tools enable the study of signaling at a level of detail and granularity that matches our understanding of these systems.**

***Keywords* — signaling pathways, rule-based modeling, computational methods, systems biology.**

I. Background

Two basic challenges constrain any approach to mathematical modeling of biological systems. First, signaling systems are combinatorially complex because proteins undergo multiple post-translational modifications and form multiple non-covalent complexes [1]. For example, the tail of the EGF receptor (EGFR) has at least 12 tyrosine phosphorylation sites that bind effector proteins [2], where the phosphorylation state of these sites alone (ignoring binding) leads to 4096 (2^{12}) unique states. When also considering homodimerization of EGFR, there are more than eight million possible species, and this number continues to increase dramatically when factoring in the binding of different EGFR-effectors. Unfortunately, many modeling methods, such as systems of differential equations, require an explicit *a priori* enumeration of all possible species, forcing modelers to drastically (and largely arbitrarily) curb combinatorial complexity at the outset. The assumptions made to reduce the combinatorial complexity and facilitate simulation can ultimately bias model behavior.

Second, empirical information about signaling pathways is subject to frequent revision. Models must be able to easily adapt to an evolving body of knowledge and assumptions about system components. The formal elements in many model methods do not correspond directly to individual pieces of knowledge, meaning that not only are these models difficult to read and understand, but often these models must be completely rewritten to incorporate new knowledge or change a modeling assumption.

II. Results

To address these issues, we have developed a rule-based modeling framework that avoids the need to explicitly specify all possible species in the model. The framework is built on a formal, context-free semantic language called Kappa that we created to represent proteins and reactions in a compact yet intuitive manner [3,4]. The granularity of the rule-based Kappa description of reactions closely mirrors the granularity of natural language descriptions of events (and thus the underlying knowledge itself), enabling models built in Kappa to be easily modified and extended. We created a stochastic simulator that, like Kappa, avoids combinatorial complexity, allowing for the simulation of systems that could not be handled by other currently available tools. To leverage the power and speed of deterministic simulation methods, we also developed an algorithm that creates a greatly reduced set of differential equation whose time-evolution exactly reproduces the time-evolution of the full combinatorial complex reaction-based system.

III. Conclusion

Together, this suite of tools enables us to build and simulate models of cell signaling systems in unparalleled detail. The ability to investigate complex molecular signatures and disease states at a mechanistic level will ultimately lead to better diagnostics and more effective therapeutic treatments.

References

[1] Hlavacek WS, Faeder JR, Blinov ML, Posner RG, et al. (2006) Rules for modeling signal-transduction systems. *Science STKE* **344**, re6

[2] Schulze WX, Deng L, Mann M (2005) Phosphotyrosine interactome of the ErbB-receptor kinase family. *Mol Syst Biol* **1**, 2005.0008

[3] Danos V, Laneve C (2004) Formal molecular biology. *Theoretical Comp Sci* **325**, 69-110

[4] Danos V, Feret J, Fontana W, Harmer R, Krivine J. (2007) Rule-based modeling of cellular systems. *Lecture Notes in Comp Sci* **4703**, 17-41

[1]Plectix BioSystems, Somerville MA 02144
[2]E-mail: ty@plectix.com