Fluxes and fluctuations in biochemical models

Maarleveld, T.R.

2015

document version
Publisher's PDF, also known as Version of record

Summary

Biology, computer science, and mathematics, as independent sciences, deal with completely different aspects. Biology studies living organisms. Computer science deals with the construction, programming, operation, and use of computers. Mathematics is the study of number, quantity, and space which can be applied to different disciplines. Computational systems biology is a fast growing interdisciplinary discipline that uses computer science and (applied) mathematics to study complex biological systems.

Building a mathematical model is an attractive and effective way of using computational systems biology to study the emergent properties of a complex biological system. A computer simulation uses such a mathematical model to predict its behavior in a specific condition. These kind of simulations require following a set of mathematical operations—an algorithm. The used models and algorithms are often so complex that appropriate software is required for this purpose.

I underlined mathematical model, computer simulation, algorithm, and software for a reason. In Chapter 1 I explained that they are the key players I used to study *fluxes and fluctuations in biochemical models* (the title of this dissertation). More specifically, I used two distinct simulation approaches to study fluxes and fluctuations in biochemical models: stoichiometric simulations to study fluxes (Chapters 2–5) and stochastic simulations to study fluctuations (Chapters 6–8). Briefly, stoichiometric simulations are deterministic (the outcome is precisely determined without any room of variation). They are called stoichiometric simulations because kinetic information is ignored. In contrast, stochastic simulations take into account this kinetic information and they posses inherent randomness (the outcome of each simulation is different). This means that both simulation approaches are completely different; both approaches require also different models. This dissertation is therefore logically divided into different parts, a part about stoichiometric models and simulations and a part about stochastic models and simulations.

The part about fluxes in biochemical models starts with Chapter 2. Here we gave an overview of the state-of-the-art stoichiometric modeling and simulation approaches. We additionally discussed the biological implications of these approaches. Stoichiometric modeling and simulation is mainly used to study metabolism. This requires genome-scale stoichiometric models that cover the total metabolic potential encoded in the genome. We developed such a genome-scale stoichiometric metabolic model of the cyanobacterium *Synechocystis* (Chapter 3). There is a lot of interest in these cyanobacteria because they have

the potential to be used as a“cellular factory“ for the conversion of solar energy and carbon dioxide into various industrial products such as biofuels.

Simulating these genome-scale models often results in big data sets. The interpretation of these big data sets a challenge. To be of any biological value, interpretation of these big data sets is required in terms of biological routes; these routes can be compared with highways. Road maps and navigation systems are widely used to travel from starting point to destination. To simplify the interpretation of the data sets generated by simulating genome-scale models, we developed a hand-drawn metabolic map (Chapter 3). Our metabolic map can be read and modified by machines and interpreted by researchers. Therefore, this metabolic map can be used, for example, for the direct visualization and analysis of the (big) data sets that are generated with stoichiometric simulations.

Genome-scale stoichiometric models are often used to determine the optimal biomass yield. Perhaps surprisingly, there are typically multiple biological routes (and combinations thereof) with an optimal biomass yield. This gives rise to the optimal solution space. In Chapter 4 we extensively studied the characteristics of this space. We are the first that uniquely characterized the entire optimal solution space in terms of a specific set of metabolic flux routes. We can quickly characterize this set of metabolic flux routes with the TimoTimo (this name is derived from the navigation systems manufacturer TomTom). The TimoTimo uses a divide-and-conquer approach (Chapter 5) to quickly characterize the optimal solution space in terms of this specific set of metabolic flux routes.

Stoichiometric simulations are, as mentioned earlier, deterministic. They ignore that biological systems are inherently stochastic. While this stochasticity is often negligible in the macroscopic world because of the law of large numbers, experimental evidence indicates that significant stochastic fluctuations are present and essential. This is especially true when molecules (e.g. mRNA or transcription factors) are present at low copy numbers. There exist mathematical methods that capture these stochastic fluctuations, which are called stochastic simulation algorithms.

The application and development of these algorithms to study fluctuations in biochemical models is the topic of the second part of my dissertation. This part starts with Chapter 6 where we provided an overview of the state-of-the-art stochastic simulation algorithms. We also used several illustrative examples to demonstrate the advantage of stochastic simulations compared to deterministic simulations. For instance, stochastic simulations allow for better estimation of the kinetic model parameters.

To study fluctuations in biochemical models we developed StochPy (Chapter 7). StochPy is a comprehensive and user-friendly software package for stochastic simulations. In Chapter 7 we used various case studies to demonstrate the power of StochPy. The current state-of-the-art stochastic simulation algorithms ignore the, often important, stochastic contributions of both cell growth and division. In Chapter 8 we developed a generic stochastic simulation algorithm that takes into account the stochastic contributions of net molecule

synthesis as the stochastic contributions of cell growth and division. This novel algorithm is now implemented in StochPy and is in exact agreement with theory, and in good agreement with time-lapse microscopy data of growing micro-organisms.

In a concluding discussion (Chapter 9), I looked critically at the use and future of the key players (model, simulation, algorithm, and software) in systems biology. I expect that they are going to be a core research component to study complex biological systems. This discussion ends with my philosophy about the requirements for useful scientific software.