Computational Programming for Product Designing in Synthetic Biology

Divya Sindhu, Saurabh Sindhu
Lecturer (Resource Person), Department of Computer Science, CRM Jat College, Hisar, Haryana, India

ABSTRACT: Biological sciences pose a unique set of engineering challenges resulting from an incomplete understanding of natural biological systems and tools for manipulating them. To address these challenges, synthetic biology is advancing from developing proof-of-concept designs to focusing on core platforms for rational and high-throughput biological engineering. These platforms span the entire biological design cycle, including DNA construction, gene libraries, computational design tools and interfaces for manipulating and probing synthetic circuits in designing and construction of synthetic genetic products. Computational methods would improve the predictability of respective mechanisms and the analysis along with the interpretation of both qualitative and quantitative models on generalizable techniques in addition to application-specific designs. This article reviews the progress and challenges in synthetic biology, computational methodology, techniques and models involved in product designing and the components used as tools, softwares and databases with their applications.

KEYWORDS: Computational analysis, Synthetic biology, Biological data base, Modeling, Softwares, Synthetic products.

1. INTRODUCTION

Biological system is composed of large number of living cells, which contain nucleic acids (DNA and RNA), proteins, carbohydrates, lipids, vitamins and minerals. The fundamental basic unit of these components in the cell is chemicals and all these components have a specific degree of freedom, conformation and orientation. Living cells of organisms operate as highly complex biological computational systems able to dynamically interrogate and respond to their environment [1,2]. Synthetic biology is the engineering of biology and leads to the synthesis of complex, biologically based systems which display functions that do not exist in nature. It involves the construction and manipulation of the biological systems from the minute molecule (individual functional unit) to the functional cellular level [3]. Synthetic biology is of two types; one using unnatural molecules to reproduce natural behavior and the other interchanges the parts from one system to another to ultimately assembled system resulting in an unnatural function.

Rapid advances in DNA synthesis techniques have made it possible to engineer viruses, available bacterial genomes and various metabolic pathways [4]. Based on novel genome based methods, synthetic biology is a rigorous engineering discipline to create, control and program cellular behavior. The vast increase of DNA assembly techniques and the genetic tools currently available for synthetic biologists have been recently reviewed allowing the achievement of new functions and the production of helpful metabolites in living cells in a controlled manner [5]. In order to link computational tools to cell-free systems, synthetic biology endeavours to build artificial biological systems through the combination of molecular biology and engineering approaches [1].

Computation biology involves theoretical methods, mathematical modeling and computation simulation techniques to study the biological systems [6,7]. This discipline shares the principles, concepts, tools, features and objectives of computer science, applied mathematics, statistics, biochemistry, biophysics, molecular biology, genetics and genomics. It deals with the analysis, investigation, estimation and elucidation of dynamic interactions between genes and proteins in naturally occurring systems [8]. In this perceptive, the end goal is to finally engineer a system/organism to perform how we want it to perform [9]. The modelling component of synthetic biology allows the designing of biological circuits and the analysis of its expected behaviour. The experimental component merges models with real systems by providing quantitative data and sets of available biological units that can be used to construct circuits. Sufficient progress has been made in the combined use of modeling and experimental methods, which reinforces the idea of being
able to use engineered microbes as a technological platform [10]. But to view cells as true ‘programmable’ entities, it is now essential to develop effective strategies for assembling procedures and modules into complex, customizable larger scale systems and the steps from modules to systems represents the second wave of synthetic biology (Fig. 1). Therefore, the ability to create such systems address to innovative approaches for a wide range of applications such as bioremediation using biosensors, sustainable energy production and biomedical therapies [11].

II. COMPUTATIONAL TOOLS, TECHNIQUES AND METHODOLOGIES

The development of an integrated and extensible biological design cycle would enable biological engineers to develop high-level conceptual designs, translate these designs into potential circuit implementations using libraries of well-characterized devices, model and verify their behavior in silico, construct the designs in an automated or high-throughput fashion, and modulate and probe the resulting constructs for proper operation [12]. Continuous feedback between multiple stages in the design cycle will be necessary to enhance their performance and integration. Synthetic biologists have laid the foundational rules of biological design [13], constructed a catalog of standardized genetic parts and assembled simple circuits such as oscillators [14], switches [15] and Boolean logic gates [16]. Moreover, experimentally validated computational tools have been created for both in vivo and in vitro systems with the aim of building biological components. Synthetic biologists are making attempts to bridge the gap between the understanding of complex biological networks and main biochemical processes by comparing modelling algorithms for both systems. Lewis et al. [17] have recently proposed a framework for synthetic biologists to build novel artificial cellular systems and to identify research areas for computational model development. As the field of biology becomes increasingly quantitative, in vitro reactions remain a powerful tool for biologists and the plasticity of cell-free systems to test model predictions under minimal conditions is of relevance.

In recent years, cell-free synthetic biology is also emerging as a powerful technology aimed to understand, connect and increase the capacity of natural biological systems without using intact cells. Cell-free systems bypass cell walls and remove genetic regulation to enable direct access to the inner machinery of the cell. Therefore, the design of in vivo systems brought to the rapid development of engineering foundations for cell-free systems and offered a versatile testbed for understanding why nature’s designs work efficiently and also for enabling biosynthetic routes to novel chemicals and new classes of materials. The emergence of cell-free systems opened the way to the synthesis of novel products by application of biochemical engineering or novel bioproduction strategies, that until now have been unfeasible to produce by other means [18, 19].

Synthetic biologists have documented modelling and simulation of genetic regulatory systems as well as outlined the basic features of synthetic biology as a new engineering discipline [20]. The understanding of biological systems has been simplified by constructing biochemical pathways and constructing computational models to reproduce the behaviour of those pathways. Molecular modeling and computational methods are employed to study the natural biological system, their kinetics and dynamics. Mathematical models became more commonly integrated into the study of biology as the mode for describing biological processes. Using algorithm, statistics and biological datas, these methods could be used for studying synthetic biological products (Fig. 2) and several tools have emerged for the recreation of in vivo synthetic biological systems [21]. The multi-scale simulation of synthetic biology can be achieved by using the techniques/methods such as slow discrete region (SDR), slow continuous region (SCR), fast discrete region (FDR), fast continuous stochastic region (FCSR), fast continuous deterministic region (FCDR) and hybrid models [12,22].

A model was presented by Chalmeau et al. [23], which described the kinetics of adsorption of the expressed protein on the phospholipid bilayer (execution of genetic programs like the introduction of a positive feedback loop into a LacI-dependent gene expression system (deals with lactose sugar utilization) in lipid vesicles, producing a cell-like system that senses and responds to an external signal with a high signal-to-noise ratio [24] and other processes associated with sensing and responding to the environment such as synthetic riboswitches can be used to control protein expression under fully defined conditions in vitro and in vesicles as well as in the artificial cells [25, 26]. The effects of molecular phenomena such as encapsulation, molecular crowding, and reaction volumes on the performance of in vitro transcription and translation could provide insights into key molecular phenomena capable to impact on in vivo gene
expression [17]. These factors profoundly affect stochastic variation of gene expression, which in turn impacts the choice between mass-action and ordinary differential equations for prediction of protein synthesis (Fig. 2).

III. SOFTWARES AND DATABASES USED IN SYNTHETIC BIOLOGY

The software bridges the gap between the kinds of instructions biological designers would like to use for designing a synthetically biological compound. The software used for designing, computation and analysis of synthetic biology are known as the Synthetic Biology Software Suite (SynBioSS). This software is used for the generation, storing, retrieval and quantitative simulation of synthetic biological networks. It is open source software which can run on Windows, Mac OS, and Linux/UNIX [27-29].

Figure 1. A pathway is modeled using bacterial cell and computational tools for synthesis of product.

Public access databases such as KEGG [30], MetaCyc [31], ChEBI [32] and RHEA [33] were found useful for the design of metabolic pathways. A database containing molecular and biochemical data of enzymes, BRENDA can be useful to select the core pathway capable to produce the metabolite of interest [34]. Web servers, such as From-
Metabolite-To-Metabolite (FMM) [35] and Metabolic Route Search and Design (MRSD) [36] can also be used for designing synthetic and unique metabolic pathways in cell-free systems. Based on enzymatic efficiency and maximum pathway yields, XTMS platform can help to rank pathways [37]. To calculate the relative contribution of each enzymatic step in the pathway when optimization of particular objective function is required, Flux Balance Analysis (FBA) is commonly used [38]. This flux is based on the stoichiometry of the metabolic pathway and several computational tools are available to solve FBA [39] such as COBRA toolbox for MATLAB [40]. Metabolic Tinker can be used to identify and rank thermodynamically favourable pathways between two compounds [41], which may include novel and non-natural pathways.

IV. DESIGNING OF SYNTHETIC BIOLOGICAL PRODUCT

Synthetic biological products can be produced from bacterial species, yeast and mammalian cells. Moreover, the properties related with *in vitro* systems have been subjugated to create complex circuitry in cell-free systems. Designing of synthetically biological compound through computational approach uses Bottom-Up method and circuit engineering analogy. Thus, it is applicable for both simpler and complex organisms [42].

![Diagram of computational methodologies used for designing of synthetic biological product and its analysis.](image-url)

**Figure 2.** Computational methodologies used for designing of synthetic biological product and its analysis.

For a well-established synthetic model, building block materials such as oscillator, switches, logical gates (AND) and comparators are required. Besides these components algorithm is also used. In case of genetic product/molecule, the
required components are promoter sequence, operator sequence, ribosome binding site, termination site, reporter protein, activator protein and repressor protein. One of the most common examples of phenotypic synthetic biological construct is of tetracycline which explains the transcription and translation process of protein synthesis both in presence and absence of tetracycline i.e. in off state and in on state [43-44].

Thus, designing include the construction of an in vitro bistable circuit from synthetic transcriptional switches, the synthetic in vitro transcriptional oscillators and the bottom-up construction of in vitro switchable memories. In processes using biochemical circuits, artificial transcriptional networks consisting of synthetic DNA switches, regulated by RNA signals acting as transcription repressors and two enzymes, a T7 RNA polymerase and an Escherichia coli ribonuclease H, were engineered. Construction of larger synthetic circuits provides an opportunity for evaluating model prediction and design of complex biochemical systems that could be used to control nanoscale devices and artificial cells [45]. The negative feedback loop in the system produces oscillation in the activities of the promoters. Synthetic transcriptional oscillators could prove valuable for systematic exploration of biochemical circuit design principles [46].

V. MODELS USED FOR PRODUCT DESIGNING IN SYNTHETIC BIOLOGY

Synthetic biologists are making attempts to bridge the gap between the understanding of complex biological networks and main biochemical processes by comparing modelling algorithms for both systems. Framework has been proposed for synthetic biologists to build novel artificial cellular systems and to identify research areas for computational model development [17]. There is a broad kind of models available for synthetic biologists. Some areas of potential growth have been identified for researchers interested in developing tools for cell-free systems. The difficulty of modelling in vivo systems stems from the context-dependency of reaction parameters. A wide diversity of equations describing the behaviour of synthetic biological systems and parameters of these equations are generally unknown. Among databases for obtaining enzymatic reaction constants, KEGG [30], BRENDA [34], and ExPASy [47] can be mentioned.

BioNumbers have also collected measurements of biological systems [48] and have been used in the modeling of a yeast-bacteria ecosystem, in a predictor of antimicrobial protein efficacy and in a computational representation of distributive metabolic networks [49-50]. Based on the level of resolution and time scale, various models are used in designing of synthetic biology.

(i). Deterministic models

These models usually consist of differential equations which predict the kinetics of a biological network based on its initial conditions and on past dynamics of the system [51]. They have been used to reproduce synthetic gene networks including inverters, switches, band-pass filters, multi-cellular networks and oscillators, and have also been applied to simulate the behaviour of tumor-invading bacteria, prokaryotic circuits capable of producing artificial analog computation and a transcriptional oscillator [52]. These models make use of Michaelis-Menten equations to describe each chemical reaction.

(ii). Stochastic modeling

Stochastic models of cellular processes can be formulated following the master equations. The variation caused by incomplete distribution of reactants within a system is the extrinsic noise, whereas intrinsic noise is the variation caused by the discrete nature of small-scale chemical reactions [53]. A deep impact on biological systems was observed due to both classes of noises during replication, during variation observed at small reaction volumes within a cell and by bursts of translation caused by limited transcriptional activity [54-55]. Stochastic models have been applied to understand sporulation dynamics of B. subtilis, adaptation of a genetic circuit in response to divergent environmental conditions and control of a bacterial population composition with a gene circuit [56-58].

(iii). Exploratory models

This type of model is used to guide the design of biological circuits. Exploratory models can also be used to analyze the effect of intrinsic noise, extrinsic noise and variation of kinetic parameters on synthetic genetic machinery [57, 59]. These exploratory models have been used to design in vivo pathways, such as a Boolean network of transcriptional switches implemented in yeast, a multiplexor circuit in E. coli and an inducible system of fluorescent reporters in
mammalian cells [60-62]. To imitate strategies, they changed from the engineering disciplines to automated biological design, combining known modules into more complicated architectures. Firstly, an automated design algorithm registered a library of biochemical parts with defined kinetic parameters and interactions described as ordinary differential equations [63]. Then, the algorithm selected certain parts and arranged them into motifs [60, 61, 64]. Transcriptional networks in vivo and in vitro have the same circuit architecture and basic components. The parameters used by these automatic genetic design programs are actually determined in vitro, which would make the assembly of in vitro circuits more accurate than in vivo circuits.

(iv). Molecular dynamic models

All-atom (AA) and coarse-grained (CG) models are categorized by the level of detail. The degree of detail addressed by each algorithm is determined by force fields, consisting of a set of mathematical functions and parameters that describe interactions between molecules that construct these models. Different force fields such as CHARMM, GROMOS, AMBER, and MARTINI can be described [65, 66]. The selection of AA or CG model depends on the context of the research study. If the detailed atom-atom interactions are not required, then, CG models are appropriate and the computational cost must be taken into account. However, as computational hardware and software continue to progress, it is possible to use AA models to describe dynamics over a longer time range [67]. AA-models are useful tools in lipid membrane simulation, where every atom of the solute and solvent in the system is explicitly simulated.

(v). Hybrid models

In all the above mentioned models, atomistic scale information is obtained and further transformed to lower resolution representations to achieve simulations at larger time- and/or length-scales [68]. A framework has been developed attempting to incorporate physical and chemical methods to simulate cellular functions [69]. Interestingly, experts predict that computational modelling of interactions between lipid membranes and transcription/translation machinery will provide unique insights into robustness of gene expression and enhance their capacity to control artificial cells [17, 22].

VI. CONCLUSION

With the advances in technology, research interests are shifting towards computational approaches for designing of genetic materials like DNA, RNA and proteins with newer biological behaviour and it may be used to build logical and informative architecture which is essential for vitality of life [2,20]. Synthetic biology is applicable for synthetic gene regulatory network and possesses strong predictive character. Mathematical modeling of biological science is possible through computational methods, can be written in algorithm form and rational designing and engineering is possible. Molecular kinetic and dynamic parameters can be determined through simulation approach. This will be beneficial for betterment in biotechnological and medical fields. In future, it would be possible to study the behaviour of higher level biological functions in silico and modelling can generate design principles for rational biological engineering. Synthetic biology could be applied in development of biosensor technology, reflecting on the features that make it a useful tool for designing and constructing engineered biological systems for sensing application [70]. Another application could be to use as therapeutic delivery tools, discussing their potential to advance the specificity, efficiency and safety of the current generation of cell and gene therapies, including how they can be used to confer curative effects improving those of conventional therapeutics [71]. A new perspective for the combinatorial biosynthesis of natural products that could reinvigorate drug discovery by using synthetic biology in combination with synthetic chemistry was also described [72]. Approaches for using computational modeling of synthetic biology perturbations to analyze endogenous biological circuits have been developed, with a particular focus on signaling and metabolic pathways. For accurately quantification of the expression of genes of interest in synthetic and systems biotechnology, a quantitative method based on flow cytometry and a super-folder green fluorescent protein was developed to act at single-cell resolution in Streptomyces. This work will facilitate the functional optimization of gene clusters and the drug discovery process in these organisms [73]. On the other hand, searching for alternative strategies as antibiotic therapies become obsolete due to bacterial resistance, mathematical models and simulations guide the development of complex technologies such as communication and transportation systems. In this sense, models that guide the development of
new antibiotic technologies span multiple molecular and cellular scales and facilitate the development of a novel technology [74]. Rational engineering has yielded cellular devices able to produce potent anti-malarial compounds, to detect and kill pathogenic bacteria and cancer cells and to reprogram cell fate [2, 75]. Advancements in DNA synthesis and assembly have enabled the rapid development of higher-order genetic circuits [76] of medical relevance [77]. In particular, approaches that are both practical as well as built on solid statistical principles are selected for application in systems biology [78]. Finally, computer simulation allows researchers to build a common framework for designing biological networks. In vitro reactions remain a powerful tool for experimental biologists and as the field of biology becomes ever more quantitative, it is important to take advantage of the plasticity of cell-free systems to test model predictions under simplified and minimal conditions.

REFERENCES

International Journal of Innovative Research in Science, Engineering and Technology
(An ISO 3297: 2007 Certified Organization)
Vol. 5, Issue 5, May 2016


International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016