Machine Learning Optimization of Evolvable Artificial Cells

Filippo Caschera  
FLinT  
Institute of Physics and Chemistry  
University of Southern Denmark  
+45 6550 4438  
filippo@ifk.sdu.dk

Steen Rasmussen  
FLinT  
Institute of Physics and Chemistry  
University of Southern Denmark  
+45 6550 4438  
steen@ifk.sdu.dk

Martin M Hanczyc  
FLinT  
Institute of Physics and Chemistry  
University of Southern Denmark  
+45 6550 4438  
martin@ifk.sdu.dk

ABSTRACT

An evolvable artificial cell is a chemical or biological complex system assembled in laboratory. The system is rationally designed to show life-like properties. In order to achieve an optimal design for the emergence of minimal life, a systematic search over possible ingredient combinations can be explored. A machine learning approach (Evo-DoE) could be applied to explore this experimental space and define optimal interactions according to a specific fitness function. Herein an implementation of an evolutionary design of experiments to optimize chemical and biochemical systems based on a machine learning process is presented. The optimization proceeds over generations of experiments in an iterative loop until optimal compositions are discovered. The fitness function is experimentally measured every time the loop is closed. Two examples of complex systems, namely a liposomal drug formulation and an in vitro cell-free expression system are presented as examples of optimization of molecular interactions in high dimensional space of compositions. These represent, for instance, the modules or subsystems that could be optimized by “mixing the protocols” to achieve the high level of sophistication that artificial cells require. In addition a replication cycle of oil in water emulsions is presented. They represent the container for the artificial cells.

KEYWORDS

Machine learning, experimental design, drug design, cell-free expression system, artificial cells, evolutionary programming.

1. INTRODUCTION

The optimization of a liposomal drug formulation and the protein synthesis of a cell-free expression system based on a machine learning process (Evo-DoE) are demonstrations that complex systems can be engineered to obtain targeted properties. The experiments are conducted in iterative cycle, exploiting a neural network type algorithm, and the fitness function value is calculated every time the loop is closed. To start the optimization process, the experimental space is sparsely sampled with a random selection of experiments. Successively the models of the desired response for the experimental data are built followed by sparse sampling of the experimental space, and then the process repeats [1].

2. RESULTS

2.1 Optimization of lipid membrane composition

A lipid vesicle as the container for the artificial cell mimics some properties of the biological membranes. The minimal cell may have a great potential of technological innovation [2]. In this section the results of optimization of a liposomal drug formulation with a machine learning process are presented. The figure 2 shows the fitness of recipes found by Evo-DoE during all generations of experiments. The system was quickly optimized after individually testing 450 individual recipes from a space hundred of times larger. The ability of intercalating an amphiphilic drug (Amphotericin) into the bilayers of phospholipids vesicles was measured as output to build the fitness function.

REFERENCES


ACKNOWLEDGMENTS

The teams of scientists at Protolife Inc. and at the FLinT Center at the University of Southern Denmark are gratefully acknowledged for contributing to the development of this research topic.

2.2 Optimization of cell-free expression system for in vitro protein synthesis

The cell-free expression system is a commercial E. Coli cell extract with defined sets of components used to express proteins inside the aqueous core of vesicles from DNA [3]. The graph shown in figure 2, represents the experimentally measured evolutionary progress of Evo-DoE. The fitness function was defined as the maximum in fluorescence measured at different time intervals during the expression of the green fluorescence protein (GFP). As a result a 300 % improvement in protein yield was measured, compared to a benchmark recipe, was measured. Evo-DoE identified the optimal ingredient mixture in the designed experimental space.

Figure 2: Experimentally measured fitness over eight generations. The standard is shown in blue and randomly chosen recipes in red. The green represents the combinations chosen from Evo-DoE.

Figure 3: Photos of oil droplets replication cycle in the lab. The droplets have dyes of two different colors.

3. ICT

3.1 Cell – Scope as ICT Interface

Two examples of experimental chemical systems optimization based on machine learning algorithms are presented. The high-throughput experiments were conducted with a robotic workstation for liquid handling. The combinations tested during the screening were indicated by the predictive algorithm (Evo-DoE), which was able to improve the fitness functions over generations of experiments. For example, to optimize the laboratory replication cycle for the oil in water emulsions, we could envision that different machine learning approaches be engaged by groups in different locations. This could be done inexpensively by engineering an ICT device as a cell-scene [8]. This would provide a mobile phone platform, which integrated with imaging analysis software and machine learning algorithms running elsewhere, which would be able to analyze and control and optimize the experimental system. The remote control can be done through an automatic process, where robotic workstations are used. The oil droplets can be used for the co-localization of the artificial cells components and since the replication-cycle, shown in Figure 4 is iterative; evolution could be a parameter that is measured over time.

Figure 1: Rank order of all tested formulations found with Evo-DoE. The black bar is the standard recipe.