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 high dimensional space of possible experimental 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 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 requires. 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 from 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 B) into the bilayers of phospholipids vesicles was measured as output to build the fitness function.

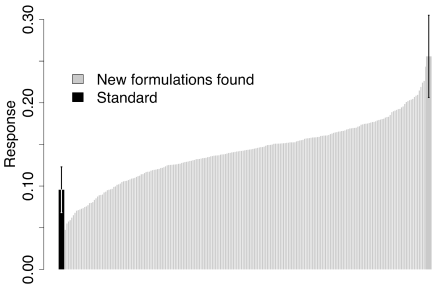

Figure 1: Rank order of all tested formulations found with Evo-DoE. The black bar is the standard recipe

The experiments were conducted in high-throughput screening and the fitness function values were measured with a spectrophotometric assay, which measured directly the amount of drug complexed in the lipid mixture.

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 indentified the optimal ingredient mixture in the designed experimental space

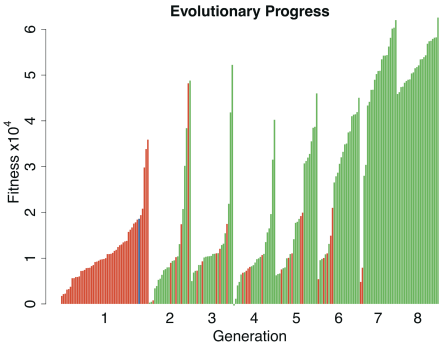

Figure 2: Experimentally measured fitness over eight generations.. The standard is shown in blue and randomly chosen recipes in red. The green represents the combinations chosen from Evo-DoE

2.3 Replication cycle of an oil droplets system

The components are compartmentalized in order to achieve the emergence of minimal life [4]. The artificial entities could be programmed to show evolution and thereby selection could be applied during their life cycle [5]. The compartments can be based on lipid vesicles (Bioreactor) or on oil in water emulsions (Oil droplet) [6, 7]. The up-take of resources needed for the metabolism and evolution can be obtained exploiting compartments dynamics, in particular fusion and fission. The first mechanism can be used for the turnover of the building blocks constituting the complex system and the fission can be exploited to apply artificial selection. The physical-chemical instabilities of the oil droplets are exploited to induce fusion and fission. The replication cycle presented in figure 3 is iterative, and the system dynamic properties can be easily controlled in the laboratory. Exploiting these mechanisms a life-cycle of the oil in water emulsion compartments is developed, but not yet optimized with machine learning.

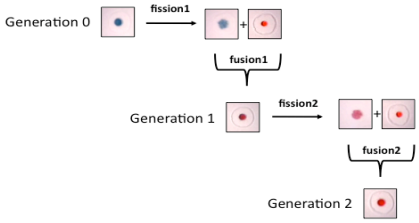

Figure 3: Photos of oil droplets replication cycle in the lab. The droplets have dyes of two different colors.

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 interface as a cell-scope [8]. This would provide a mobile phone platform, which integrated with imaging analysis software and 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.

ACKNOWLEDGMENTS

The teams of scientists at ProtoLife Inc. and at the FLinT Center at the University of Southern Denmark are gratefully acknow-ledged for contributing to the development of this research topic.

REFERENCES

1) Caschera F., Gazzola G., Bedau M.A., Bosch Moreno C., Buchanan A., et al. Automated Discovery of Novel Drug Formulations Using Predictive Iterated High Throughput Experimentation. *PLoS ONE* 5(1): e8546. 2010
2) Pohrille A. and Deamer D. Artificial cells prospects for biotechnology. *TRENDS in Biotechnology* 20:123-128. 2002
3) Noireaux V., Libchaber A. A. vesicle bioreactor as a step toward an artificial cell assembly. *Proc. Natl. Acad. Sci.* USA 101:17669–17674. 2004
4) Rasmussen S., Chen L., Deamer D., Krauker D.C., Packard N.H., Stadler P.F., Bedau M.A. Evolution. Transitions from nonliving to living matter. *Science* 303: 963-5. 2004
5) Szostak J.W., Bartel D.P., Luisi P.L. Synthesizing life. *Nature* 409:387–390. 2001
6) Ichihashi N., Matsuura T., Kita H., Sunami T., Suzuki H., Yomo T. Constructing Partial Models of Cells. *Cold. Spring. Harb. Perspect. Biol.* June; 2(6): a004945. 2010
7) Hanczyc M.M., Toyota T., Ikegami T., Packard N., Sugawara T. 2007. Fatty acid chemistry at the oil-water interface: self-propelled oil droplets. *JACS* 129: 9386-91. 2007
8) Breslaurer D.N., Maamari R.N., Switz N.A., Lam W.A., Fletcher D.A. Mobile phone based clinical microscopy for global health applications. PLoS ONE 4(7): e6320. 2009