

Title: A Tool-Chain to Accelerate Synthetic Biological Engineering

Authors: Jacob Beal, Ron Weiss, Douglas Densmore, Aaron Adler,  
Jonathan Babb, Swapnil Bhatia, Noah Davidson,  
Traci Haddock, Fusun Yaman, Richard Schantz, Joseph Loyall

Work sponsored by DARPA I20 under contract HR0011-10-C-0168; the views and conclusions contained in this document are those of the authors and not DARPA or the U.S. Government.

There is a pressing need for design automation tools for synthetic biological systems. Compared to electronic circuits, cellular information processing has more complex elementary components with a much greater complexity of interactions between components. Moreover, chemical computation within a cell is strongly affected both by other computations taking place within the cell and by the cell's native metabolic processes and its external environment. All this adds up to a design flow that is currently highly iterative, error-prone, and extremely slow: critical problems that must all be addressed in order to realize the potential of synthetic biology.

We have been developing a tool-chain approach to decomposing the problems of design and assembly into automatable fragments. Practitioners using our tool-chain will be able to design organisms using higher level descriptions, which are automatically transformed into genetic regulatory network designs, then assembled into DNA samples ready for in vivo execution. At the same time, the tool-chain is free and open software that will allow researchers to incorporate their own design tools, thereby disseminating their results to the community and enhancing the capabilities of the tool-chain.

The current prototype begins, at the top end, with the Proto spatial computing language (Beal & Bachrach, 2006), which has a dataflow model of parallelism that maps well onto the continuous parallel flow of information through a genetic regulatory network (Beal & Bachrach, 2008). The biological compiler extension of Proto uses a motif-based compilation system to translate high-level programs into an abstract design for a genetic regulatory network for a boolean transcriptional logic circuit (Beal et al., in review). Preliminary results with this system show that it can translate compact representations of complex information processing programs into genetic regulatory network designs that perform correctly in ODE chemical simulation using a Hill equation model.

These designs are then translated into DNA parts by the MatchMaker system (Yaman et al., in review), which is integrated with Clotho, an

extensible framework for biological data integration and exchange (Densmore et al., '10). MatchMaker maps the design onto a set of available DNA parts using a constraint-based search for an equivalent interaction subgraph. The set of implementing DNA parts is then serialized into a final design, which is then fed to an assembly planner and will be able to be constructed from laboratory samples of DNA with the aid of a laboratory automation robot.

Preliminary end-to-end results from our prototype tool-chain are promising: the designs generated are equivalent to ones already being constructed in the laboratory to implement the same functionality. These results show the promise of a tool-chain approach to design automation for synthetic biology systems: design may be potentially accelerated by orders of magnitude through early fault detection (by simulation or verification analysis at each stage of the tool-chain), reduction of human error (by abstracting away details and routinizing more aspects of design), and application of AI techniques to automate exploration of the design space.