Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Structural Biology

Synthetic biology through biomolecular design and engineering

Kevin Channon[1], Elizabeth HC Bromley[1] and Derek N Woolfson[1,2]

Synthetic biology is a rapidly growing field that has emerged in a global, multidisciplinary effort among biologists, chemists, engineers, physicists, and mathematicians. Broadly, the field has two complementary goals: To improve understanding of biological systems through mimicry and to produce bio-orthogonal systems with new functions. Here we review the area specifically with reference to the concept of synthetic biology space, that is, a hierarchy of components for, and approaches to generating new synthetic and functional systems to test, advance, and apply our understanding of biological systems. In keeping with this issue of *Current Opinion in Structural Biology*, we focus largely on the design and engineering of biomolecule-based components and systems.

Addresses
[1] School of Chemistry, University of Bristol, Bristol BS8 1TS, UK
[2] Department of Biochemistry, University of Bristol, Bristol BS8 1TD, UK

Corresponding author: Woolfson, Derek N (D.N.Woolfson@bristol.ac.uk)

Current Opinion in Structural Biology 2008, **18**:491–498

This review comes from a themed issue on
Engineering and design
Edited by Dek Woolfson and Lynne Regan

Available online 5th August 2008

0959-440X/$ – see front matter

DOI 10.1016/j.sbi.2008.06.006

Introduction

Complexity in Nature is astounding, and attempts to mimic it present considerable challenges and potential rewards. This complexity stems from a hierarchical organization of biomolecular components and layers of interactions between them. Encouragingly, many aspects of both the components and their interactions are becoming increasingly understood. As outlined below, this hierarchical view and our improved understanding of and ability to engineer biology are the cornerstones of synthetic biology.

We refer to the potentially vast arena in which synthetic biologists can operate as *synthetic biology space*. This is represented as a plot of complexity of components against some indicator of how divergent from Nature these are (i.e. how 'synthetic' the components are) (Figure 1). We find this useful in two respects: First it provides a framework to chart routes toward the common goal of creating multi-component, encapsulated, functional systems; second, it allows a wide variety of studies to be grouped into a small number of general approaches. We believe that this will be useful in defining and, hopefully, helping to develop the exciting and broad area of synthetic biology.

For this review, because of the breadth of topics that contribute to this emerging field, we found it necessary to refer to classic studies from the past two decades, reviews in various areas from the past five years, as well as more recent work from the past three years.

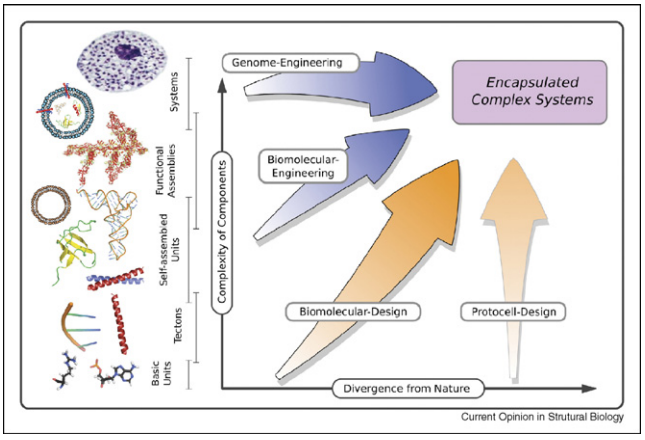
Synthetic biology space: hierarchies of components, interactions and approaches

At the base of the hierarchy is a set of *basic units*—amino acids, nucleic acids, sugars and lipids[1] (Figure 1). One level of complexity above these are what might be termed *tectons*. This term is borrowed from supramolecular chemistry [1], where it is used to describe programmed molecular components and nanoscale building blocks [2]. An example of a nucleic acid tecton would be a short oligonucleotide containing the information for further assembly through interactions with other tectons. Similarly, an amino acid based tecton would be a polypeptide designed to form stretches of self-assembling *α-helix* or *β-strands*. Importantly, a tecton is something more than a simple element of secondary structure: It implies that the element contains information about its further assembly into prescribed higher order structures. Combining tectons leads to the next level in the hierarchy, in which *self-assembled units* are formed through interactions programmed into tectons. For peptides and proteins, autonomous folding motifs would be self-assembling units. By prudent organization of such units one can arrive at *functional assemblies*. As with tectons, the definitions of *self-assembling unit* and *functional assembly* encompass functional protein and DNA tertiary and quaternary structures. With further organization, interacting networks of functional assemblies – that is, *systems* – can be constructed. In Nature, complex interacting components of a system are almost always contained, or encapsulated, within lipid membranes, which enable cells to maintain control over their environments, and the biochemical processes they conduct.

Here, we use the concept of *synthetic biology space* (Figure 1) to capture the assembly of these various components and

[1] There are, of course, many other small molecules involved in biological processes. However, a large fraction of the structural complexity of organisms can be represented in this small subset of building blocks.

Figure 1



Synthetic biology space [2]. An approach or study in synthetic biology is resolved according to where it (or its natural equivalent) appears in the natural hierarchy (*y* axis), and by some measure of how synthetic it is (*x* axis). Colored arrows indicate approximate routes through synthetic biology space taken by studies in any of the four approaches to synthetic biology described in the main text: genome engineering, biomolecular engineering, biomolecular design, and artificial protocell design. Blue arrows indicate approaches usually conducted *in vivo* and orange arrows indicate *in vitro* approaches. On the left, the various levels in the natural hierarchy and their ranges are described, along with illustrative natural examples from various points in the hierarchy.

the resulting hierarchy, as well as to highlight approaches and recent studies under the general umbrella of synthetic biology. At first, some of the studies appear to have entirely different origins and objectives. However, by placing them in a common framework, it is possible to recognize how each contributes toward the shared goal of generating complex synthetic systems.

Synthetic biologists may access the hierarchy of Nature at any level by making alterations to existing natural systems at one, or a number of levels. Studies in synthetic biology can generally be classed as either *in vivo* or *in vitro* [3] and may be further subdivided according to the approach they take to the problem at hand: *genome engineering*, *biomolecular engineering*, *biomolecular-design* or *protocell-design projects* (Figure 1) [2]. These are not sharp-edged definitions, but broad classifications that we find useful. For example, genome engineering refers to approaches like that taken by Venter and colleagues to construct synthetic chromosomes for whole or minimal organisms; biomolecular engineering includes approaches such as the BioBricks initiative, which aims to create a toolkit of functional units (usually natural protein components) that can be introduced to present new orthogonal functions in living cells; the biomolecular-design approach refers to the general idea of the *de novo* design and combination of biomolecular components; and the protocell approach includes ambitious projects to make self-replicating, encapsulated systems from entirely synthetic components.

The task of each approach is similar: To create a more synthetic entry at a higher level of complexity by manipulating a part of the preceding level. Hence, advances in synthetic biology tend to take us up and to the right in the synthetic biology space of Figure 1. The most complex and least natural systems – which are likely to be the most difficult to achieve – tend to be found in the top-right portion of synthetic biology space. Indeed, this region currently remains unoccupied, and we speculate that this will be reached only by using a combination of basis sets[2] contained by a membrane (or membrane-like) layer.

Routes to complex systems

Basic units

One challenge is to increase the number of building blocks, and by implication the repertoire of chemistries, that are accessible to synthetic biology. For example, bio-inspired building blocks include the DNA analog PNA, in which purine and pyrimidine base-pairing is natural, but the sugar–phosphate backbone is replaced by *N*-(2-aminoethyl)-glycine units linked by peptide bonds [4]. Various non-proteinogenic α-amino acids can also be used to bring new chemistry to proteins within the more conventional polypeptide framework [5]. Further along the divergence axis (Figure 1) β-peptides have been shown to adopt distinct secondary structures [6] making them potential tectons (Figure 2) and recently tertiary structures bringing them to the self-assembled unit level [7•,8•]. This has been extended yet further up the hierarchy with the recent assembly of β-peptide-based fibers [9].

It is also possible to design building blocks that diverge completely from Nature. Examples include pyridine dicarboxamides [10]; helicogenic polyisocyantide systems [11]; anthranilamides [12]; quinoline oligamides [13]; benzoylurea oligomers [14] (Figure 2); and naphthyridine foldamers [15]. Onto these relatively rigid templates, a variety of functional side groups can be appended, analogous to natural side-chain functionality. This 'blank-sheet' approach is very tempting, as it presents the possibility that synthetic systems could effectively isolate structure from function, and hence makes nanostructure design much more routine. However, there are inevitably drawbacks: First, there is no guarantee that such systems would reproduce the complexity and dynamics inherent in, and so important to natural assemblies; second, for purely synthetic building blocks there is currently no analogous infrastructure for their replication, regulation, segregation, and turnover, as there is for the natural and, in some cases, non-natural [5,16,17] building blocks of cells.

[2] This could be an 'all-natural' basis – mixtures of proteins, nucleic acids, sugars, and lipids, among others – or an entirely synthetic basis, or indeed anywhere in between.