

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Synthetic Biology & Bioinformatics Prospects in the Cancer Arena

Lígia R. Rodrigues and Leon D. Kluskens

*IBB – Institute for Biotechnology and Bioengineering, Centre of Biological Engineering,  
University of Minho, Campus de Gualtar, Braga  
Portugal*

## 1. Introduction

Cancer is the second leading cause of mortality worldwide, with an expected 1.5-3.0 million new cases and 0.5-2.0 million deaths in 2011 for the US and Europe, respectively (Jemal et al., 2011). Hence, this is an enormously important health risk, and progress leading to enhanced survival is a global priority. Strategies that have been pursued over the years include the search for new biomarkers, drugs or treatments (Rodrigues et al., 2007). Synthetic biology together with bioinformatics represents a powerful tool towards the discovery of novel biomarkers and the design of new biosensors.

Traditionally, the majority of new drugs has been generated from compounds derived from natural products (Neumann & Neumann-Staubitz, 2010). However, advances in genome sequencing together with possible manipulation of biosynthetic pathways, constitute important resources for screening and designing new drugs (Carothers et al. 2009). Furthermore, the development of rational approaches through the use of bioinformatics for data integration will enable the understanding of mechanisms underlying the anti-cancer effect of such drugs (Leonard et al., 2008; Rocha et al., 2010).

Besides in biomarker development and the production of novel drugs, synthetic biology can also play a crucial role in the level of specific drug targeting. Cells can be engineered to recognize specific targets or conditions in our bodies that are not naturally recognized by the immune system (Forbes, 2010).

Synthetic biology is the use of engineering principles to create, in a rational and systematic way, functional systems based on the molecular machines and regulatory circuits of living organisms or to re-design and fabricate existing biological systems (Benner & Sismour, 2005). The focus is often on ways of taking parts of natural biological systems, characterizing and simplifying them, and using them as a component of a highly unnatural, engineered, biological system (Endy, 2005). Virtually, through synthetic biology, solutions for the unmet needs of humankind can be achieved, namely in the field of drug discovery. Indeed, synthetic biology tools enable the elucidation of disease mechanisms, identification of potential targets, discovery of new chemotherapeutics or design of novel drugs, as well as the design of biological elements that recognize and target cancer cells. Furthermore, through synthetic biology it is possible to develop economically attractive microbial production processes for complex natural products.

Bioinformatics is used in drug target identification and validation, and in the development of biomarkers and tools to maximize the therapeutic benefit of drugs. Now that data on cellular signalling pathways are available, integrated computational and experimental projects are being developed, with the goal of enabling *in silico* pharmacology by linking the genome, transcriptome and proteome to cellular pathophysiology. Furthermore, sophisticated computational tools are being developed that enable the modelling and design of new biological systems. A key component of any synthetic biology effort is the use of quantitative models (Arkin, 2001). These models and their corresponding simulations enable optimization of a system design, as well as guiding their subsequent analysis. Dynamic models of gene regulatory and reaction networks are essential for the characterization of artificial and synthetic systems (Rocha et al., 2008). Several software tools and standards have been developed in order to facilitate model exchange and reuse (Rocha et al., 2010).

In this chapter, synthetic biology approaches for cancer diagnosis and drug development will be reviewed. Specifically, examples on the design of RNA-based biosensors, bacteria and virus as anti-cancer agents, and engineered microbial cell factories for the production of drugs, will be presented.

## 2. Synthetic biology: tools to design, build and optimize biological processes

Synthetic biology uses biological insights combined with engineering principles to design and build new biological functions and complex artificial systems that do not occur in Nature (Andrianantoandro et al., 2006). The building blocks used in synthetic biology are the components of molecular biology processes: promoter sequences, operator sequences, ribosome binding sites (RBS), termination sites, reporter proteins, and transcription factors. Examples of such building blocks are given in Table 1.

Great developments of DNA synthesis technologies have opened new perspectives for the design of very large and complex circuits (Purnick & Weiss, 2009), making it now affordable to synthesize a given gene instead of cloning it. It is possible to synthesize *de novo* a small virus (Mueller et al., 2009), to replace the genome of one bacterium by another (Lartigue et al., 2007) and to make large chunks of DNA coding for elaborate genetic circuits. Software tools to simulate large networks and the entire panel of omics technologies to analyze the engineered microorganism are available (for details see section 3). Finally, the repositories of biological parts (e.g. Registry of Standard Biological Parts (<http://partsregistry.org/>)) will increase in complexity, number and reliability of circuits available for different species.

Currently, the design and synthesis of biological systems are not decoupled. For example, the construction of metabolic pathways or any circuit from genetic parts first requires a collection of well characterized parts, which do not yet fully exist. Nevertheless, this limitation is being addressed through the development and compilation of standard biological parts (Kelly et al., 2009). When designing individual biological parts, the base-by-base content of that part (promoter, RBS, protein coding region, terminator, among others) is explicitly dictated (McArthur IV & Fong, 2010). Rules and guidelines for designing genetic parts at this level are being established (Canton et al., 2008). Particularly, an important issue when designing protein-coding parts is codon optimization, encoding the same amino acid sequence with an alternative, preferred nucleotide sequence. Although a particular sequence, when expressed, may be theoretically functional, its expression may be far from optimal or even completely suppressed due to codon usage bias in the heterologous host.

Genetic part	Examples	Rationale
<i>Transcriptional control</i>		
Constitutive promoters	<i>lacIq, SV40, T7, sp6</i>	“Always on” transcription
Regulatory regions	<i>tetO, lacO, ara, gal4, rhl</i> box	Repressor and activator sites
Inducible promoters	<i>ara, ethanol, lac, gal, rhl, lux, fdhH, sal, glnK, cyc1</i>	Control of the promoter by induction or by cell state
Cell fate regulators	GATA factors	Control cell differentiation
<i>Transcriptional control</i>		
RNA interference (RNAi)	Logic functions, RNAi repressor	Genetic switch, logic evaluation and gene silencing
Riboregulators	Ligand-controlled ribozymes	Switches for detection and actuation
Ribosome binding site	Kozak consensus sequence mutants	Control the level of translation
<i>Post-transcriptional control</i>		
Phosphorylation cascades	Yeast phosphorylation pathway	Modulate genetic circuit behavior
Protein receptor design	TNT, ACT and EST receptors	Control detection thresholds and combinatorial protein function
Protein degradation	Ssra tags, peptides rich in Pro, Glu, Ser and Thr	Protein degradation at varying rates
Localization signals	Nuclear localization, nuclear export and mitochondrial localization signals	Import or export from nucleus and mitochondria
<i>Others</i>		
Reporter genes	GFP, YFP, CFP, LacZ	Detection of expression
Antibiotic resistance	ampicilin, chloramphenicol	Selection of cells

Table 1. Genetic elements used as components of synthetic regulatory networks (adapted from McArthur IV & Fong, 2010 and Purnick & Weiss, 2009). Legend: CFP, cyan fluorescent protein; GFP, green fluorescent protein; YFP, yellow fluorescent protein.

Codon optimization of coding sequences can be achieved using freely available algorithms such as Gene Designer (see section 3). Besides codon optimization, compliance with standard assembly requirements and part-specific objectives including activity or specificity modifications should be considered. For example, the BioBrick methodology requires that parts exclude four standard restriction enzyme sites, which are reserved for use in assembly (Shetty et al., 2008). Extensive collections of parts can be generated by using a naturally occurring part as a template and rationally modifying it to create a library of that particular genetic part. Significant progress in this area has been recently demonstrated for promoters and RBS (Ellis et al., 2009; Salis et al., 2009). Ellis and co-workers (2009) constructed two promoter libraries that can be used to tune network behavior *a priori* by fitting mathematical promoter models with measured parameters. By using this model-guided design approach the authors were able to limit the variability of the system and increase predictability.

However, it is well-known that noisy or leaky promoters can complicate the system design. In these cases a finer control over expression can be established by weakening the binding strength of the downstream gene (Ham et al., 2006), or by using two promoter inputs to drive transcription of an output via a modular AND gate (Anderson et al., 2006). Additionally, modular and scalable RNA-based devices (aptamers, ribozymes, and transmitter sequences) can be engineered to regulate gene transcription or translation (Win & Smolke, 2007).

Design at the pathway level is not only concerned with including the necessary parts, but also with controlling the expressed functionality of those parts. Parts-based synthetic metabolic pathways will require tunable control, just as their natural counterparts which often employ feedback and feed-forward motifs to achieve complex regulation (Purnick & Weiss, 2009). Using a synthetic biology approach, the design of DNA sequences encoding metabolic pathways (e.g. operons) should be relatively straightforward. Synthetic scaffolds and well-characterized families of regulatory parts have emerged as powerful tools for engineering metabolism by providing rational methodologies for coordinating control of multigene expression, as well as decoupling pathway design from construction (Ellis et al., 2009). Pathway design should not overlook the fact that exogenous pathways interact with native cellular components and have their own specific energy requirements. Therefore, modifying endogenous gene expression may be necessary in addition to balancing cofactor fluxes and installing membrane transporters (Park et al., 2008).

After designing parts, circuits or pathways, the genomic constructs ought to be manufactured through DNA synthesis. Nucleotide's sequence information can be outsourced to synthesis companies (e.g. DNA2.0, GENEART or Genscript, among others). The convenience of this approach over traditional cloning allows for the systematic generation of genetic part variants such as promoter libraries. Also, it provides a way to eliminate restriction sites or undesirable RNA secondary structures, and to perform codon optimization. The ability to make large changes to DNA molecules has resulted in standardized methods for assembling basic genetic parts into larger composite devices, which facilitate part-sharing and faster system-level construction, as demonstrated by the BioBrick methodology (Shetty et al., 2008) and the Gateway cloning system (Hartley, 2003). Other approaches based on type II restriction enzymes, such as Golden Gate Shuffling, provide ways to assemble many more components together in one step (Engler et al., 2009). A similar one-step assembly approach, circular polymerase extension cloning (CPEC), avoids the need for restriction-ligation, or single-stranded homologous recombination altogether (Quan & Tian, 2009). Not only is this useful for cloning single genes, but also for assembling parts into larger sequences encoding entire metabolic pathways and for generating combinatorial part libraries. On a chromosomal level, disruption of genes in *Escherichia coli* and other microorganisms has become much faster with the development of RecBCD and lambda RED-assisted recombination systems (Datsenko & Wanner, 2000), allowing the insertion, deletion or modification by simply using linear gene fragments. Additionally, multiplex automated genome engineering (MAGE) has been introduced as another scalable, combinatorial method for producing large-scale genomic diversity (Wang et al., 2009). This approach makes chromosomal modification easier by simultaneously mutating target sites across the chromosome. Plasmid-based expression and chromosomal integration are the two common vehicles for implementing synthetic metabolic pathways. Recently, the chemically inducible chromosomal evolution (CICHE) was proposed as a long-



term expression alternative method (Tyo et al., 2009). This new method avoids complications associated with plasmid replication and segregation, and can be used to integrate multiple copies of genes into the genome. All these techniques will provide technical platforms for the rapid synthesis of parts and subsequent pathways.

The majority of the synthetic biology advances has been achieved purely *in vitro* (Isalan et al., 2008), or in microorganisms involving the design of small gene circuits without a direct practical application, although scientifically very exciting. These studies have offered fundamental insight into biological processes, like the role and sources of biological noise; the existence of biological modules with defined properties; the dynamics of oscillatory behavior; gene transcription and translation; or cell communication (Alon, 2003; Kobayashi et al., 2004). An interesting example of a larger system that has been redesigned is the refactoring of the T7 bacteriophage (Chan et al., 2005). Another successful example has been the production of terpenoid compounds in *E. coli* (Martin et al., 2003) and *Saccharomyces cerevisiae* (Ro et al., 2006) that can be used for the synthesis of artemisinin. Bacteria and fungi have long been used in numerous industrial microbiology applications, synthesizing important metabolites in large amounts. The production of amino acids, citric acid and enzymes are examples of other products of interest, overproduced by microorganisms. Genetic engineering of strains can contribute to the improvement of these production levels. Altogether, the ability to engineer biological systems will enable vast progress in existing applications and the development of several new possibilities. Furthermore, novel applications can be developed by coupling gene regulatory networks with biosensor modules and biological response systems. An extensive RNA-based framework has been developed for engineering ligand-controlled gene regulatory systems, called ribozyme switches. These switches exhibit tunable regulation, design modularity, and target specificity and could be used, for example, to regulate cell growth (Win & Smolke, 2007). Engineering interactions between programmed bacteria and mammalian cells will lead to exciting medical applications (Anderson et al., 2006). Synthetic biology will change the paradigm of the traditional approaches used to treat diseases by developing “smart” therapies where the therapeutic agent can perform computation and logic operations and make complex decisions (Andrianantoandro et al., 2006). There are also promising applications in the field of living vectors for gene therapy and chemical factories (Forbes, 2010; Leonard et al., 2008).

### **3. Bioinformatics: a rational path towards biological behavior predictability**

In order to evolve as an engineering discipline, synthetic biology cannot rely on endless trial and error methods driven by verbal description of biomolecular interaction networks. Genome projects identify the components of gene networks in biological organisms, gene after gene, and DNA microarray experiments discover the network connections (Arkin, 2001). However, these data cannot adequately explain biomolecular phenomena or enable rational engineering of dynamic gene expression regulation. The challenge is then to reduce the amount and complexity of biological data into concise theoretical formulations with predictive ability, ultimately associating synthetic DNA sequences to dynamic phenotypes.

#### **3.1 Models for synthetic biology**

The engineering process usually involves multiple cycles of design, optimization and revision. This is particularly evident in the process of constructing gene circuits (Marguet et

al., 2007). Due to the large number of participating species and the complexity of their interactions, it becomes difficult to intuitively predict a design behavior. Therefore, only detailed modeling can allow the investigation of dynamic gene expression in a way fit for analysis and design (Di Ventura et al., 2006). Modeling a cellular process can highlight which experiments are likely to be the most informative in testing model hypothesis, and for example allow testing for the effect of drugs (Di Bernardo et al., 2005) or mutant phenotypes (Segre et al., 2002) on cellular processes, thus paving the way for individualized medicine.

Data are the precursor to any model, and the need to organize as much experimental data as possible in a systematic manner has led to several excellent databases as summarized in Table 2. The term “model” can be used for verbal or graphical descriptions of a mechanism underlying a cellular process, or refer to a set of equations expressing in a formal and exact manner the relationships among variables that characterize the state of a biological system (Di Ventura et al., 2006). The importance of mathematical modeling has been extensively demonstrated in systems biology (You, 2004), although its utility in synthetic biology seems even more dominant (Kaznessis, 2009).

Name	Website
BIND (Biomolecular Interaction Network Database)	<a href="http://www.bind.ca/">http://www.bind.ca/</a>
Brenda (a comprehensive enzyme information system)	<a href="http://www.brenda.uni-koeln.de/">http://www.brenda.uni-koeln.de/</a>
CSNDB (Cell Signaling Networks Database)	<a href="http://geo.nihs.go.jp/csndb/">http://geo.nihs.go.jp/csndb/</a>
DIP (Database of Interacting Proteins)	<a href="http://dip.doe-mbi.ucla.edu/">http://dip.doe-mbi.ucla.edu/</a>
EcoCyc/Metacyc/BioCyc (Encyclopedia of E. coli genes and metabolism)	<a href="http://ecocyc.org/">http://ecocyc.org/</a>
EMP (Enzymes and Metabolic Pathways Database)	<a href="http://www.empproject.com/">http://www.empproject.com/</a>
GeneNet (information on gene networks)	<a href="http://wwwmgs.bionet.nsc.ru/mgs/systems/genenet/">http://wwwmgs.bionet.nsc.ru/mgs/systems/genenet/</a>
Kegg (Kyoto Encyclopedia of Genes and Genomes)	<a href="http://www.genome.ad.jp/kegg/kegg.html">http://www.genome.ad.jp/kegg/kegg.html</a>
SPAD (Signaling Pathway Database)	<a href="http://www.grt.kyushu-u.ac.jp/enydoc/">http://www.grt.kyushu-u.ac.jp/enydoc/</a>
RegulonDB ( <i>E. coli</i> K12 transcriptional network)	<a href="http://regulondb.ccg.unam.mx/">http://regulondb.ccg.unam.mx/</a>
ExPASy-beta (Bioinformatics Resource Portal)	<a href="http://beta.expasy.org/">http://beta.expasy.org/</a>

Table 2. Databases of molecular properties, interactions and pathways (adapted from Arkin, 2001).

Model-driven rational engineering of synthetic gene networks is possible at the level of topologies or at the level of molecular components. In the first one, it is considered that molecules control the concentration of other molecules, e.g. DNA-binding proteins regulate the expression of specific genes by either activation or repression. By combining simple regulatory interactions, such as negative and positive feedback and feed-forward loops, one may create more complex networks that precisely control the production of protein

molecules (e.g. bistable switches, oscillators, and filters). Experimentally, these networks can be created using existing libraries of regulatory proteins and their corresponding operator sites. Examples of these models are the oscillator described by Gardner *et al* (2000) and repressilator by Elowitz and Leibler (2000). In the second level, the kinetics and strengths of molecular interactions within the system are described. By altering the characteristics of the components, such as DNA-binding proteins and their corresponding DNA sites, one can modify the system dynamics without modifying the network topology. Experimentally, the DNA sequences that yield the desired characteristics of each component can be engineered to achieve the desired protein-protein, protein-RNA, or protein-DNA binding constants and enzymatic activities. For example, Alon and co-workers (2003) showed how simple mutations on the DNA sequence of the lactose operon can result in widely different phenotypic behavior.

Various mathematical formulations can be used to model gene circuits. At the population level, gene circuits can be modeled using ordinary differential equations (ODEs). In an ODE formulation, the dynamics of the interactions within the circuit are deterministic. That is, the ODE formulation ignores the randomness intrinsic to cellular processes, and is convenient for circuit designs that are thought to be less affected by noise or when the impact of noise is irrelevant (Marguet *et al.*, 2007). An ODE model facilitates further sophisticated analyses, such as sensitivity analysis and bifurcation analysis. Such analyses are useful to determine how quantitative or qualitative circuit behavior will be impacted by changes in circuit parameters. For instance, in designing a bistable toggle switch, bifurcation analysis was used to explore how qualitative features of the circuit may depend on reaction parameters (Gardner *et al.*, 2000). Results of the analysis were used to guide the choice of genetic components (genes, promoters and RBS) and growth conditions to favor a successful implementation of designed circuit function. However, in a single cell, the gene circuit's dynamics often involve small numbers of interacting molecules that will result in highly noisy dynamics even for expression of a single gene. For many gene circuits, the impact of such cellular noise may be critical and needs to be considered (Di Ventura *et al.*, 2006). This can be done using stochastic models (Arkin, 2001). Different rounds of simulation using a stochastic model will lead to different results each time, which presumably reflect aspects of noisy dynamics inside a cell. For synthetic biology applications, the key of such analysis is not necessarily to accurately predict the exact noise level at each time point. This is not possible even for the simplest circuits due to the "extrinsic" noise component for each circuit (Elowitz *et al.*, 2002). Rather, it is a way to determine to what extent the designed function can be maintained and, given a certain level of uncertainty or randomness, to what extent additional layers of control can minimize or exploit such variations. Independently of the model that is used, these can be evolved *in silico* to optimize designs towards a given function. As an example, genetic algorithms were used by Francois and Hakim (2004) to design gene regulatory networks exhibiting oscillations.

In most attempts to engineer gene circuits, mathematical models are often purposefully simplified to accommodate available computational power and to capture the qualitative behavior of the underlying systems. Simplification is beneficial partially due to the limited quantitative characterization of circuit elements, and partially because simpler models may better reveal key design constraints. The limitation, however, is that a simplified model may fail to capture richer dynamics intrinsic to a circuit. Synthetic models combine features of mathematical models and model organisms. In the engineering of genetic networks,



synthetic biologists start from mathematical models, which are used as the blueprints to engineer a model out of biological components that has the same materiality as model organism but is much less complex. The specific characteristics of synthetic models allow one to use them as tools in distinguishing between different mathematical models and evaluating results gained in performing experiments with model organisms (Loettgers, 2007).

### 3.2 Computational tools for synthetic biology

Computational tools are essential for synthetic biology to support the design procedure at different levels. Due to the lack of quantitative characterizations of biological parts, most design procedures are iterative requiring experimental validation to enable subsequent refinements (Canton et al., 2008). Furthermore, stochastic noise, uncertainty about the cellular environment of an engineered system, and little insulation of components complicate the design process and require corresponding models and analysis methods (Di Ventura et al., 2006). Many computational standards and tools developed in the field of systems biology (Wierling et al., 2007) are applicable for synthetic biology as well.

As previously discussed, synthetic gene circuits can be constructed from a handful of basic parts that can be described independently and assembled into interoperating modules of different complexity. For this purpose, standardization and modularity of parts at different levels is required (Canton et al., 2008). The Registry of Standard Biological Parts constitutes a reference point for current research in synthetic biology and it provides relevant information on several DNA-based synthetic or natural building blocks. Most computational tools that specifically support the design of artificial gene circuits use information from the abovementioned registry. Moreover, many of these tools share standardized formats for the input/output files. The System Biology Markup Language (SBML) (<http://sbml.org>) defines a widely accepted, XML-based format for the exchange of mathematical models in biology. It provides a concise representation of the chemical reactions embraced by a biological system. These can be translated into systems of ODEs or into reaction systems amenable to stochastic simulations (Alon, 2003). Despite its large applicability to simulations, SBML currently lacks modularity, which is not well aligned with parts registries in synthetic biology. Alternatively, synthetic gene systems can be described according to CellML language which is more modular (Cooling et al., 2008).

One important feature to enable the assembly of standard biological parts into gene circuits is that they share common inputs and outputs. Endy (2005) proposed RNA polymerases and ribosomes as the molecules that physically exchange information between parts. Their fluxes, measured in PoPS (Polymerase Per Second) and in RiPS (Ribosomes Per Second) represent biological currents (Canton et al., 2008). This picture, however, does not seem sufficient to describe all information exchanges even in simple engineered gene circuits, since other signal carriers like transcription factors and environmental “messages” should be explicitly introduced and not indirectly estimated by means of PoPS and RiPS (Marchisio & Stelling, 2008). Based on the assumption that parts share common input/output signals, several computational tools have been proposed for gene circuit design, as presented in Table 3. Comparing these circuit design tools it is obvious that we are still far from an ideal solution. The software tools differ in many aspects such as scope of parts and circuit descriptions, the mode of user interaction, and the integration with databases or other tools.

<b>Circuit design and simulation</b>	
Biojade	<a href="http://web.mit.edu/jagoler/www/biojade/">http://web.mit.edu/jagoler/www/biojade/</a>
Tinkercell	<a href="http://www.tinkercell.com/Home">http://www.tinkercell.com/Home</a>
Asmparts	<a href="http://soft.synth-bio.org/asmparts.html">http://soft.synth-bio.org/asmparts.html</a>
ProMoT	<a href="http://www.mpimagdeburg.mpg.de/projects/promot">http://www.mpimagdeburg.mpg.de/projects/promot</a>
GenoCAD	<a href="http://www.genocad.org/genocad/">http://www.genocad.org/genocad/</a>
GEC	<a href="http://research.microsoft.com/gec">http://research.microsoft.com/gec</a>
TABASCO	<a href="http://openwetware.org/wiki/TABASCO">http://openwetware.org/wiki/TABASCO</a>
<b>Circuit optimization</b>	
Genetdes	<a href="http://soft.synth-bio.org/genetdes.html">http://soft.synth-bio.org/genetdes.html</a>
RoVerGeNe	<a href="http://iasi.bu.edu/~batt/rovergene/rovergene.htm">http://iasi.bu.edu/~batt/rovergene/rovergene.htm</a>
<b>DNA and RNA design</b>	
Gene Designer	<a href="https://www.dna20.com/index.php?pageID=220">https://www.dna20.com/index.php?pageID=220</a>
GeneDesign	<a href="http://www.genedesign.org/">http://www.genedesign.org/</a>
UNAFold	<a href="http://www.bioinfo.rpi.edu/applications/hybrid/download.php">http://www.bioinfo.rpi.edu/applications/hybrid/download.php</a>
Vienna RNA package	<a href="http://www.tbi.univie.ac.at/~ivo/RNA/">http://www.tbi.univie.ac.at/~ivo/RNA/</a>
Zinc Finger Tools	<a href="http://www.scripps.edu/mb/barbas/zfdesign/zfdesignhome.php">http://www.scripps.edu/mb/barbas/zfdesign/zfdesignhome.php</a>
<b>Protein Design</b>	
Rosetta	<a href="http://www.rosettacommons.org/main.html">http://www.rosettacommons.org/main.html</a>
RAPTOR	<a href="http://www.bioinformaticssolutions.com/products/raptor/index.php">http://www.bioinformaticssolutions.com/products/raptor/index.php</a>
PFP	<a href="http://dragon.bio.purdue.edu/pfp/">http://dragon.bio.purdue.edu/pfp/</a>
Autodock 4.2	<a href="http://autodock.scripps.edu/">http://autodock.scripps.edu/</a>
HEX 5.1	<a href="http://webloria.loria.fr/~ritchied/hex/">http://webloria.loria.fr/~ritchied/hex/</a>
<b>Integrated workflows</b>	
SynBioSS	<a href="http://synbioss.sourceforge.net/">http://synbioss.sourceforge.net/</a>
Clotho	<a href="http://biocad-server.eecs.berkeley.edu/wiki/index.php/Tools">http://biocad-server.eecs.berkeley.edu/wiki/index.php/Tools</a>
Biskit	<a href="http://biskit.sf.net/">http://biskit.sf.net/</a>

Table 3. Computational design tools for synthetic biology (adapted from Marchisio & Stelling, 2009; Matsuoka et al., 2009; and Prunick & Weiss, 2009)

Biojade was one of the first tools being reported for circuit design (Goler, 2004). It provides connections to both parts databases and simulation environments, but it considers only one kind of signal carrier (RNA polymerases). It can invoke the simulator TABASCO (Kosuri et al., 2007), thus enabling genome scale simulations at single base-pair resolution. CellDesigner (Funahashi et al., 2003) has similar capabilities for graphical circuit composition. However, parts modularity and consequently circuit representation do not appear detailed enough. Another tool for which parts communicate only by means of PoPS, but not restricted to a single mathematical framework, is the Tinkercell. On the contrary, in Asmparts (Rodrigo et al., 2007a) the circuit design is less straightforward and intuitive because the tool lacks a Graphic User Interface. Nevertheless, each part exists as an independent SBML module and the model kinetics for transcription and translation permit to limit the number of parameters necessary for a qualitative system description. Marchisio and Stelling (2008) developed a new framework for the design of synthetic circuits where

each part is modeled independently following the ODE formalism. This results in a set of composable parts that communicate by fluxes of signal carriers, whose overall amount is constantly updated inside their corresponding pools. The model also considers transcription factors, chemicals and small RNAs as signal carriers. Pools are placed among parts and devices: they store free signal carriers and distribute them to the whole circuit. Hence, polymerases and ribosomes have a finite amount; this permits to estimate circuit scalability with respect to the number of parts. Mass action kinetics is fully employed and no approximations are required to depict the interactions of signal carriers with DNA and mRNA. The authors implemented the corresponding models into ProMoT (Process Modeling Tool), software for the object-oriented and modular composition of models for dynamic processes (Mirschel et al., 2009). GenoCAD (Czar et al., 2009) and GEC (Pedersen & Phillips, 2009) introduce the notions of a grammar and of a programming language for genetic circuit design, respectively. These tools use a set of rules to check the correct composition of standard parts. Relying on libraries of standard parts that are not necessarily taken from the Registry of Standard Biological Parts, these programs can translate a circuit design into a complete DNA sequence. The two tools differ in capabilities and possible connectivity to other tools.

The ultimate goal of designing a genetic circuit is that it works, i.e. that it performs a given function. For that purpose, optimization cycles to establish an appropriate structure and a good set of kinetic parameters values are required. These optimization problems are extremely complex since they involve the selection of adequate parts and appropriate continuous parameter values. Stochastic optimization methods (e.g. evolutionary algorithms) attempt to find good solutions by biased random search. They have the potential for finding globally optimal solutions, but optimization is computationally expensive. On the other hand, deterministic methods (e.g. gradient descent) are local search methods, with less computational cost, but at the expense of missing good solutions.

The optimization problem can be tackled by tools such as Genetdes (Rodrigo et al., 2007b) and OptCircuit (Dasika & Maranas, 2008). They rely on different parts characterizations and optimization algorithms. Genetdes uses a stochastic method termed "Simulated Annealing" (Kirkpatrick et al., 1983), which produces a single solution starting from a random circuit configuration. As a drawback, the algorithm is more likely to get stuck in a local minimum than an evolutionary algorithm. OptCircuit, on the contrary, treats the circuit design problem with a deterministic method (Bansal et al., 2003), implementing a procedure towards a "local" optimal solution. Each of these optimization algorithms requires a very simplified model for gene dynamics where, for instance, transcription and translation are treated as a single step process. Moreover, the current methods can cope only with rather small circuits. Another tool that has been described by Batt and co-workers (2007), RoVerGeNe, addresses the problem of parameter estimation more specifically. This tool permits to tune the performance and to estimate the robustness of a synthetic network with a known behavior and for which the topology does not require further improvement.

Detailed design of synthetic parts that reproduce the estimated circuit kinetics and dynamics is a complex task. It requires computational tools in order to achieve error free solutions in a reasonable amount of time. Other than the placement/removal of restriction sites and the insertion/deletion of longer motifs, mutations of single nucleotides may be necessary to tune part characteristics (e.g. promoter strength and affinity toward regulatory factors). Gene Designer (Villalobos et al., 2006) is a complete tool for building artificial DNA

segments and codon usage optimization. GeneDesign (Richardson et al., 2006) is another tool to design long synthetic DNA sequences. Many other tools are available for specific analysis of the DNA and RNA circuit components. The package UNAFold (Markham & Zuker, 2008) predicts the secondary structure of nucleic acid sequences to simulate their hybridizations and to estimate their melting temperature according to physical considerations. A more accurate analysis of the secondary structure of ribonucleic acids can be performed through the Vienna RNA package (Hofacker, 2003). Binding sites along a DNA chain can be located using Zinc Finger Tools (Mandell & Barbas, 2006). These tools allows one to search DNA sequences for target sites of particular zinc finger proteins (Kaiser, 2005), whose structure and composition can also be arranged. Thus, gene control by a class of proteins with either regulation or nuclease activity can be improved. Furthermore, tools that enable promoter predictions and primers design are available, such as BDGP and Primer3. Another relevant task in synthetic biology is the design and engineering of new proteins. Many tools have been proposed for structure prediction, homology modeling, function prediction, docking simulations and DNA-protein interactions evaluation. Examples include the Rosetta package (Simons et al., 1999); RAPTOR (Xu et al., 2003); PFP (Hawkins et al., 2006); Autodock 4.2 (Morris et al., 2009) and Hex 5.1 (Ritchie, 2008).

Further advance in computational synthetic biology will result from tools that combine and integrate most of the tasks discussed, starting with the choice and assembly of biological parts to the compilation and modification of the corresponding DNA sequences. Examples of such tools comprise SynBioSS (Hill et al., 2008); Clotho and Biskit (Grunberg et al., 2007). Critical elements are still lacking, such as tools for automatic information integration (literature and databases), and tools that re-use standardized model entities for optimal circuit design. Overall, providing an extended and integrated information technology infrastructure will be crucial for the development of the synthetic biology field.

#### **4. A roadmap from design to production of new drugs**

Biological systems are dynamic, that is they mutate, evolve and are subject to noise. Currently, the full knowledge on how these systems work is still limited. As previously discussed, synthetic biology approaches involve breaking down organisms into a hierarchy of composable parts, which is useful for conceptualization purposes. Reprogramming a cell involves the creation of synthetic biological components by adding, removing, or changing genes and proteins. Nevertheless, it is important to notice that assembly of parts largely depends on the cellular context (the so-called chassis), thus restraining the abstraction of biological components into devices and modules, and their use in design and engineering of new organisms or functions.

One level of abstraction from the DNA synthesis and manipulation is parts production, which optimization can be accomplished through either rational design or directed evolution. Applying rational design to parts alteration or creation is advantageous, in that it cannot only generate products with a defined function, but it can also produce biological insights into how the designed function comes about. However, it requires prior structural knowledge of the part, which is frequently unavailable. Directed evolution is an alternative method that can effectively address this limitation. Many synthetic biology applications will require parts for genetic circuits, cell-cell communication systems, and non-natural metabolic pathways that cannot be found in Nature, simply because Nature is not in need of them (Dougherty & Arnold, 2009). In essence, directed evolution begins with the generation



of a library containing many different DNA molecules, often by error-prone DNA replication, DNA shuffling or combinatorial synthesis (Crameri et al., 1998). The library is next subjected to high-throughput screening or selection methods that maintain a link between genotype and phenotype in order to enrich the molecules that produce the desired function. Directed evolution can also be applied at other levels of biological hierarchy, for example to evolve entire gene circuits (Yokobayashi et al., 2002). Rational design and directed evolution should not be viewed as opposing methods, but as alternate ways to produce and optimize parts, each with their own unique strengths and weaknesses. Directed evolution can complement this technique, by using mutagenesis and subsequent screening for improved synthetic properties (Brustad & Arnold, 2010). In addition, methods have been developed to incorporate unnatural amino acids in peptides and proteins (Voloshchuk & Montclare, 2009). This will expand the toolbox of protein parts, and add beneficial effects, such as increased *in vivo* stability, when incorporated in proteinaceous therapeutics. Also, the development of *de novo* enzymes has seen a significant increase lately. The principle of computational design uses the design of a model, capable of stabilizing the transition state of a reaction. From there on, individual amino acids are positioned around it to create a catalytic site that stabilizes the transition state. The mRNA display technique resembles phage display and is a technique for the *in vitro* selection and evolution of proteins. Translated proteins are associated with their mRNA via a puromycin linkage. Selection occurs by binding to an immobilized substrate, after which a reverse transcriptase step will reveal the cDNA and thus the nucleotide sequence (Golynskiy & Seelig, 2010). If the selection step includes measurement of product formation from the substrate, novel peptides with catalytic properties can be selected.

For the design, engineering, integration and testing of new synthetic gene networks, tools and methods derived from experimental molecular biology must be used (for details see section 2). Nevertheless, progress on these tools and methods is still not enough to guarantee the complete success of the experiment. As a result, design of synthetic biological systems has become an iterative process of modeling, construction, and experimental testing that continues until a system achieves the desired behavior (Purnick & Weiss, 2009). The process begins with the abstract design of devices, modules, or organisms, and is often guided by mathematical models (Koide et al., 2009). Afterwards, the newly constructed systems are tested experimentally. However, such initial attempts rarely yield fully functional implementations due to incomplete biological information. Rational redesign based on mathematical models improves system behavior in such situations (Koide et al., 2009; Prather & Martin, 2008). Directed evolution is a complimentary approach, which can yield novel and unexpected beneficial changes to the system (Yokobayashi et al., 2002). These retooled systems are once again tested experimentally and the process is repeated as needed. Many synthetic biological systems have been engineered successfully in this fashion because the methodology is highly tolerant to uncertainty (Matsuoka et al., 2009). Figure 1 illustrates the above mentioned iterative approach used in synthetic biology.

Since its inception, metabolic engineering aims to optimize cellular metabolism for a particular industrial process application through the use of directed genetic modifications (Tyo et al., 2007). Metabolic engineering is often seen as a cyclic process (Nielsen, 2001), where the cell factory is analyzed and an appropriate target is identified. This target is then experimentally implemented and the resulting strain is characterized experimentally and, if necessary, further analyses are conducted to identify novel targets. The application of



synthetic biology to metabolic engineering can potentially create a paradigm shift. Rather than starting with the full complement of components in a wild-type organism and piecewise modifying and streamlining its function, metabolic engineering can be attempted from a bottom-up, parts-based approach to design by carefully and rationally specifying the inclusion of each necessary component (McArthur IV & Fong, 2010). The importance of rationally designing improved or new microbial cell factories for the production of drugs has grown substantially since there is an increasing need for new or existing drugs at prices that can be affordable for low-income countries. Large-scale re-engineering of a biological circuit will require systems-level optimization that will come from a deep understanding of operational relationships among all the constituent parts of a cell. The integrated framework necessary for conducting such complex bioengineering requires the convergence of systems and synthetic biology (Koide et al., 2009). In recent years, with advances in systems biology (Kitano, 2002), there has been an increasing trend toward using mathematical and computational tools for the *in silico* design of enhanced microbial strains (Rocha et al., 2010).

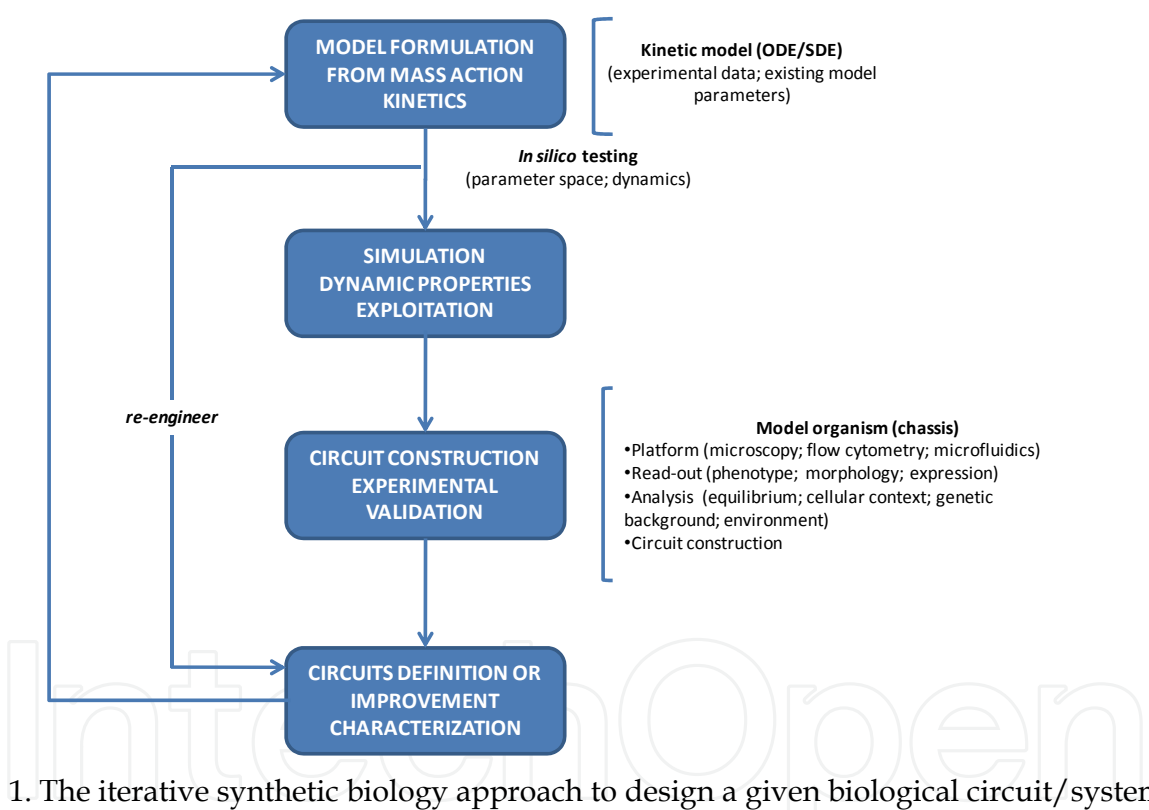


Fig. 1. The iterative synthetic biology approach to design a given biological circuit/system.

Current models in both synthetic and systems biology emphasize the relationship between environmental influences and the responses of biological networks. Nevertheless, these models operate at different scales, and to understand the new paradigm of rational systems re-engineering, synthetic and systems biology fields must join forces (Koide et al., 2009). Synthetic biology and bottom-up systems biology methods extract discrete, accurate, quantitative, kinetic and mechanistic details of regulatory sub-circuits. The models generated from these approaches provide an explicit mathematical foundation that can ultimately be used in systems redesign and re-engineering. However, these approaches are confounded by high dimensionality, non-linearity and poor prior knowledge of key dynamic parameters (Fisher & Henzinger, 2007) when scaled to large systems.

Consequently, modular sub-network characterization is performed assuming that the network is isolated from the rest of the host system. The top-down systems biology approach is based on data from high-throughput experiments that list the complete set of components within a system in a qualitative or semi-quantitative manner. Models of overall systems are similarly qualitative, tending toward algorithmic descriptions of component interactions. Such models are amenable to the experimental data used to develop them, but usually sacrifice the finer kinetic and mechanistic details of the molecular components involved (Price & Shmulevich, 2007). Bridging systems and synthetic biology approaches is being actively discussed and several solutions have been suggested (Koide et al., 2009).

A typical synthetic biology project is the design and engineering of a new biosynthetic pathway in a model organism (chassis). Generally, *E. coli* is the preferred chassis since it is well-studied, easy to manipulate, and its reproduction in biological cultures is very handy. Initially, relevant databases like Kegg and BioCyc (Table 2) can be consulted for identifying all the possible metabolic routes that allow the production of a given drug from metabolites that exist in native *E. coli*. Then, for each reaction, the species that are known to possess the corresponding enzymes/genes must be identified. This step is relevant, since most often the same enzyme exhibits different kinetic behavior among different species. Information on sequences and kinetic parameters can be extracted from the above-mentioned sources, relevant literature and also from Brenda and ExPASy databases. Afterwards, the information collected is used to build a family of dynamic models (Rocha et al., 2008, 2010) that enable the simulation of possible combinations regarding pathway configuration and the origin of the enzymes (associated with varying kinetic parameters). Optflux tool can be used for simulations and metabolic engineering purposes (<http://www.optflux.org/>). Using the same input (a fixed amount of precursor) it is possible to select the configuration that allows obtaining higher drug yields. Furthermore, through the use of genome-scale stoichiometric models coupled with the dynamic model it is possible to understand the likely limitations regarding the availability of the possible precursors. In fact, if the precursor for the given drug biosynthesis is a metabolic intermediate, possible limitations in its availability need to be addressed in order to devise strategies to cope with it. Based on this information, the next step involves the construction of the enzymatic reactions that will lead to the production of the drug from a metabolic precursor in *E. coli*. The required enzymes are then synthesized based on the gene sequences previously selected from the databases. The cloning strategy may include using a single plasmid with two different promoters; using two different plasmids, with different copy numbers and/or origins of replication; or ultimately integrating it into the genome, in order to allow fine tuning of the expression of the various enzymes necessary. Finally, a set of experiments using the engineered bacterium needs to be performed to evaluate its functionality, side-product formation and/or accumulation, production of intermediate metabolites and final product (desired drug). In the case of the previously mentioned artemisinin production, DNA microarray analysis and targeted metabolic profiling were used to optimize the synthetic pathway, reducing the accumulation of toxic intermediates (Kizer et al., 2008). These types of methodologies enable the validation of the drug production model and the design of strategies to further improve its production.

## 5. Novel strategies for cancer diagnosis and drug development

Cancer is a main issue for the modern society and according to the World Health Organization it is within the top 10 of leading causes of death in middle- and high-income

countries. Several possibilities to further improve existing therapies and diagnostics, or to develop novel alternatives that still have not been foreseen, can be drawn using synthetic biology approaches. Promising future applications include the development of RNA-based biosensors to produce a desired response *in vivo* or to be integrated in a cancer diagnosis device; the design and engineering of bacteria that can be programmed to target a tumor and release a therapeutic agent *in situ*; the use of virus as a tool for recognizing tumors or for gene therapy; and the large scale production of complex chemotherapeutic agents, among others.

### 5.1 RNA-based biosensors

Synthetic biology seeks for new biological devices and systems that regulate gene expression and metabolite pathways. Many components of a living cell possess the ability to carry genetic information, such as DNA, RNA, proteins, among others. RNA has a critical role in several functions (genetic translation, protein synthesis, signal recognition of particles) due to its functional versatility from genetic blueprint (e.g. mRNA, RNA virus genomes. Its catalytic function as enzyme (e.g. ribozymes, rRNA) and regulator of gene expression (e.g. miRNA, siRNA) makes it stand out among other biopolymers with a more specialized scope (e.g. DNA, proteins) (Dawid et al., 2009). Therefore, non-coding RNA molecules enable the formation of complex structures that can interact with DNA, other RNA molecules, proteins and other small molecules (Isaacs et al., 2006).

Natural biological systems contain transcription factors and regulators, as well as several RNA-based mechanisms for regulating gene expression (Saito & Inoue, 2009). A number of studies have been conducted on the use of RNA components in the construction of synthetic biologic devices (Topp & Gallivan, 2007; Win & Smolke, 2007). The interaction of RNA with proteins, metabolites and other nucleic acids is affected by the relationship between sequence, structure and function. This is what makes the RNA molecule so attractive and malleable to engineering complex and programmable functions.

#### 5.1.1 Riboswitches

One of the most promising elements are the riboswitches, genetic control elements that allow small molecules to regulate gene expression. They are structured elements typically found in the 5'-untranslated regions of mRNA that recognize small molecules and respond by altering their three-dimensional structure. This, in turn, affects transcription elongation, translation initiation, or other steps of the process that lead to protein production (Beisel & Smolke, 2009; Winkler & Breaker, 2005). Biological cells can modulate gene expression in response to physical and chemical variations in the environment allowing them to control their metabolism and preventing the waste of energy expenditure or inappropriate physiological responses (Garst & Batey, 2009). There are currently at least twenty classes of riboswitches that recognize a wide range of ligands, including purine nucleobases (purine riboswitch), amino acids (lysine riboswitch), vitamin cofactors (cobalamin riboswitch), amino sugars, metal ions (mgtA riboswitch) and second messenger molecules (cyclic di-GMP riboswitch) (Beisel & Smolke, 2009). Riboswitches are typically composed of two distinct domains: a metabolite receptor known as the aptamer domain, and an expression platform whose secondary structure signals the regulatory response. Embedded within the aptamer domain is the switching sequence, a sequence shared between the aptamer domain and the expression platform (Garst & Batey, 2009). The aptamer domain is part of the RNA

and forms precise three-dimensional structures. It is considered a structured nucleotide pocket belonging to the riboswitch, in the 5'-UTR, which when bound regulates downstream gene expression (Isaacs et al., 2006). Aptamers specifically recognize their corresponding target molecule, the ligand, within the complex group of other metabolites, with the appropriate affinity, such as dyes, biomarkers, proteins, peptides, aromatic small molecules, antibiotics and other biomolecules. Both the nucleotide sequence and the secondary structure of each aptamer remain highly conserved (Winkler & Breaker, 2005). Therefore, aptamer domains are the operators of the riboswitches.

A strategy for finding new aptamer sequences is the use of SELEX (Systemic Evolution of Ligands by Exponential enrichment method). SELEX is a combinatorial chemistry technique for producing oligonucleotides of either single-stranded DNA or RNA that specifically bind to one or more target ligands (Stoltenburg et al., 2007). The process begins with the synthesis of a very large oligonucleotide library consisting of randomly generated sequences of fixed length flanked by constant 5' and 3' ends that serve as primers. The sequences in the library are exposed to the target ligand and those that do not bind the target are removed, usually by affinity chromatography. The bound sequences are eluted and amplified by PCR to prepare for subsequent rounds of selection in which the stringency of the elution conditions is increased to identify the tightest-binding sequences (Stoltenburg et al., 2007). SELEX has been used to evolve aptamers of extremely high binding affinity to a variety of target ligands. Clinical uses of the technique are suggested by aptamers that bind tumor markers (Ferreira et al., 2006). The aptamer sequence must then be placed near to the RBS of the reporter gene, and inserted into *E. coli* (chassis), using a DNA carrier (i.e. plasmid), in order to exert its regulatory function.

Synthetic riboswitches represent a powerful tool for the design of biological sensors that can, for example, detect cancer cells, or the microenvironment of a tumor, and in the presence of a given molecule perform a desired function, like the expression *in situ* of a therapeutic agent. Several cancer biomarkers have been identified in the last decade; therefore there are many opportunities of taking these compounds as templates to design adequate riboswitches for their recognition. Alternatively, the engineering goal might be the detection of some of these biomarkers in biological samples using biosensors with aptamers as the biological recognition element, hence making it a less invasive approach. The development of aptamer-based electrochemical biosensors has made the detection of small and macromolecular analytes easier, faster, and more suited for early detection of protein biomarkers (Hianik & Wang, 2009). Multi-sensor arrays that provide global information on complex samples (e.g. biological samples) have deserved much interest recently. Coupling an aptamer to these devices will increase its specificity and selectivity towards the selected target(s). The selected target may be any serum biomarker that when detected in high amounts in biological samples can be suggestive of tumor activity.

## 5.2 Bacteria as anti-cancer agents

Bacteria possess unique features that make them powerful candidates for treating cancer in ways that are unattainable by conventional methods. The moderate success of conventional methods, such as chemotherapy and radiation, is related to its toxicity to normal tissue and inability to destroy all cancer cells. Many bacteria have been reported to specifically target tumors, actively penetrate tissue, be easily detected and/or induce a controlled cytotoxicity. The possibility of engineering interactions between programmed bacteria and mammalian



cells opens unforeseen progresses in the medical field. Emerging applications include the design of bacteria to produce therapeutic agents (*in vitro* or *in vivo*) and the use of live bacteria as targeted delivery systems (Forbes, 2010; Pawelek et al., 2003). An impressive example of these applications is described by Anderson and co-workers (2006). The authors have successfully engineered *E. coli* harboring designed plasmids to invade cancer cells in an environmentally controlled way, namely in a density-dependent manner under anaerobic growth conditions and arabinose induction. Plasmids were built containing the *inv* gene from *Yersinia pseudotuberculosis* under control of the Lux promoter, the hypoxia-responsive *fdhF* promoter, and the arabinose-inducible *araBAD* promoter. This is significant because the tumor environment is often hypoxic and allows for high bacterial cell densities due to depressed immune function in the tumor. Therefore, this work demonstrated, as a “proof of concept”, that one can potentially use engineered bacteria to target diseased cells without significantly impacting healthy cells.

Ideally, an engineered bacterium for cancer therapy would specifically target tumors enabling the use of more toxic molecules without systemic effects; be self-propelled enabling its penetration into tumor regions that are inaccessible to passive therapies; be responsive to external signals enabling the precise control of location and timing of cytotoxicity; be able to sense the local environment allowing the development of responsive therapies that can make decisions about where and when drugs are administered; and be externally detectable, thus providing information about the state of the tumor, the success of localization and the efficacy of treatment (Forbes, 2010). Indeed some of these features naturally exist in some bacteria, e.g. many genera of bacteria have been shown to preferentially accumulate in tumors, including *Salmonella*, *Escherichia*, *Clostridium* and *Bifidobacterium*. Moreover, bacteria have motility (flagella) that enable tissue penetration and chemotactic receptors that direct chemotaxis towards molecular signals in the tumor microenvironment. Selective cytotoxicity can be engineered by transfection with genes for therapeutic molecules, including toxins, cytokines, tumor antigens and apoptosis-inducing factors. External control can be achieved using gene promoter strategies that respond to small molecules, heat or radiation. Bacteria can be detected using light, magnetic resonance imaging and positron emission tomography. At last, genetic manipulation of bacteria is easy, thus enabling the development of treatment strategies, such as expression of anti-tumor proteins and including vectors to infect cancer cells (Pawelek et al., 2003). To date, many different bacterial strategies have been implemented in animal models (e.g. *Salmonella* has been tested for breast, colon, hepatocellular, melanoma, neuroblastoma, pancreatic and prostate cancer), and also some human trials (e.g. *C. butyricum* M-55 has been tested for squamous cell carcinoma, metastatic, malignant neuroma and melanoma) have been carried out (Forbes, 2010).

Ultrasound is one of the techniques often used to treat solid tumors (e.g. breast cancer); however, this technique is not always successful, as sometimes it just heats the tumor without destroying it. Therefore, we are currently engineering the heat shock response machinery from *E. coli* to trigger the release of a therapeutic agent *in situ* concurrent with ultrasound treatment. For that purpose, several modeling and engineering steps are being implemented. The strategy being pursued is particularly useful for drugs that require *in situ* synthesis because of a poor bioavailability, thereby avoiding repetitive oral doses to achieve sufficient concentration inside the cells. The use of live bacteria for therapeutic purposes naturally poses some issues (Pawelek et al., 2003), but currently the goal is to achieve the



proof-of-concept that an engineered system will enable the production of a cancer-fighting drug triggered by a temperature increase.

### 5.3 Alternative nanosized drug carriers

The design of novel tumor targeted multifunctional particles is another extremely interesting and innovative approach that makes use of the synthetic biology principles. The modest success of the traditional strategies for cancer treatment has driven research towards the development of new approaches underpinned by mechanistic understanding of cancer progression and targeted delivery of rational combination therapy.

#### 5.3.1 Viral drug delivery systems

The use of viruses, in the form of vaccines, has been common practice ever since its first use to combat smallpox. Recently, genetic engineering has enlarged the applications of viruses, since it allows the removal of pathogen genes encoding virulence factors that are present in the virus coat. As a result, it can elicit immunity without causing serious health effects in humans. In the light of gene therapy, the use of virus-based entities hold a promising future, since by nature, they are being delivered to human target cells, and can be easily manipulated genetically. As such, they may be applied to target and lyse specific cancer cells, delivering therapeutics *in situ*. Bacteriophages are viruses that specifically and only infect bacteria. They have gained more attention the last decades, mainly in phage display technology. In anti-cancer therapy, this technique has contributed enormously to the identification of new tumor-targeting molecules (Brown, 2010). *In vivo* phage display technology identified a peptide exhibiting high affinity to hepatocellular carcinoma cells (Du et al., 2010). In a different approach, a phage display-selected ligand targeting breast cancer cells was incorporated in liposomes containing siRNA. The delivered liposomes were shown to significantly downregulate the *PRDM14* gene in the MCF7 target cells (Bedi et al., 2011). In addition, the option to directly use bacteriophages as drug-delivery platforms has been explored. A recent study described the use of genetically modified phages able to target tumor cell receptors via specific antibodies resulting in endocytosis, intracellular degradation, and drug release (Bar et al., 2008). Using phage display a variety of cancer cell binding and internalizing ligands have been selected (Gao et al., 2003). Bacteriophages can also be applied to establish an immune response. Eriksson and co-workers (2009) showed that a tumor-specific M13 bacteriophage induced regression of melanoma target cells, involving tumor-associated macrophages and being Toll-like receptor-dependent. Finally, marker molecules or drugs can be chemically conjugated onto the phage surface, making it a versatile imaging or therapy vehicle that may reduce costs and improve life quality (Steinmetz, 2010). An M13 phage containing cancer cell-targeting motifs on the surface was chemically modified to conjugate with fluorescent molecules, resulting in both binding and imaging of human KB cancer cells (Li et al., 2010). Besides being genetically part of the virus, anti-tumor compounds can also be covalently linked to it. We are currently, by phage display, selecting phages that adhere and penetrate tumor cells. Following this selection, we will chemically conjugate anti-cancer compounds (e.g. doxorubicin) to bacteriophages, equipped with the cancer cell-recognizing peptides on the phage surface. We anticipate that such a multifunctional nanoparticle, targeted to the tumor using a tumor “homing” peptide, will enable a significant improvement over existing anti-cancer approaches.

#### 5.4 Microbial cell factories for the production of drugs

On a different perspective, as exemplified in section 4, synthetic biology approaches can be used for the large scale production of compounds with pharmaceutical applications. One of the easily employable approaches to develop synthetic pathways is to combine genes from different organisms, and design a new set of metabolic pathways to produce various natural and unnatural products. The host organism provides precursors from its own metabolism, which are subsequently converted to the desired product through the expression of the heterologous genes (see section 4). Existing examples of synthetic metabolic networks make use of transcriptional and translational control elements to regulate the expression of enzymes that synthesize and breakdown metabolites. In these systems, metabolite concentration acts as an input for other control elements (Andrianantoandro et al., 2006). An entire metabolic pathway from *S. cerevisiae*, the mevalonate isoprenoid pathway for synthesizing isopentyl pyrophosphate, was successfully transplanted into *E. coli*. In combination with an inserted synthetic amorpha-4, 11-diene synthase, this pathway produced large amounts of a precursor to the anti-malarial drug artemisinin. This new producing strain is very useful since a significant decrease in the drug production time and costs could be achieved (Martin et al., 2003). In addition to engineering pathways that produce synthetic metabolites, artificial circuits can be engineered using metabolic pathways connected to regulatory proteins and transcriptional control elements (Andrianantoandro et al., 2006). One study describes such a circuit based on controlling gene expression through acetate metabolism for cell-cell communication (Bulter et al., 2004). Metabolic networks may embody more complex motifs, such as an oscillatory network. A recently constructed metabolic network used glycolytic flux to generate oscillations through the signaling metabolite acetyl phosphate (Fung et al., 2005). The system integrates transcriptional regulation with metabolism to produce oscillations that are not correlated with the cell division cycle. The general concerns of constructing transcriptional and protein interaction-based modules, such as kinetic matching and optimization of reactions for a new environment, apply for metabolic networks as well. In addition, the appropriate metabolic precursors must be present. For this purpose, it may be necessary to include other enzymes or metabolic pathways that synthesize precursors for the metabolite required in a synthetic network (Leonard et al., 2008; McArthur IV & Fong, 2010).

Many polyketides and nonribosomal peptides are being used as antibiotic, anti-tumor and immunosuppressant drugs (Neumann & Neumann-Staubitz, 2010). In order to produce them in heterologous hosts, assembly of all the necessary genes that make up the synthetic pathways is essential. The metabolic systems for the synthesis of polyketides are composed of multiple modules, in which an individual module consists of either a polyketide synthase or a nonribosomal peptide synthetase. Each module has a specific set of catalytic domains, which ultimately determine the structure of the metabolic product and thus its function. Recently, Bumpus et al. (2009) presented a proteomic strategy to identify new gene clusters for the production of polyketides and nonribosomal peptides, and their biosynthetic pathways, by adapting mass-spectrometry-based proteomics. This approach allowed identification of genes that are used in the production of the target product in a species, for which a complete genome sequence is not available. Such newly identified pathways can then be copied into a new host strain that is more suitable for producing polyketides and nonribosomal peptides at an industrial scale. This exemplifies that the sources of new pathways are not limited to species with fully sequenced genomes.

The use of synthetic biology approaches in the field of metabolic engineering opens enormous possibilities, especially toward the production of new drugs for cancer treatment. Our goal is to design and model a new biosynthetic pathway for the production of natural drugs in *E. coli*. Key to this is the specification of gene sequences encoding enzymes that catalyze each reaction in the pathway, and whose DNA sequences can be incorporated into devices that lead to functional expression of the molecules of interest (Prather & Martin, 2008). Partial pathways can be recruited from independent sources and co-localized in a single host (Kobayashi et al., 2004). Alternatively, pathways can be constructed for the production of new, non-natural products by engineering existing routes (Martin et al., 2003).

## 6. Conclusion

Despite all the scientific advances that humankind has seen over the last centuries, there are still no clear and defined solutions to diagnose and treat cancer. In this sense, the search for innovative and efficient solutions continues to drive research and investment in this field. Synthetic biology uses engineering principles to create, in a rational and systematic way, functional systems based on the molecular machines and regulatory circuits of living organisms, or to re-design and fabricate existing biological systems. Bioinformatics and newly developed computational tools play a key role in the improvement of such systems. Elucidation of disease mechanisms, identification of potential targets and biomarkers, design of biological elements for recognition and targeting of cancer cells, discovery of new chemotherapeutics or design of novel drugs and catalysts, are some of the promises of synthetic biology. Recent achievements are thrilling and promising; yet some of such innovative solutions are still far from a real application due to technical challenges and ethical issues. Nevertheless, many scientific efforts are being conducted to overcome these limitations, and undoubtedly it is expected that synthetic biology together with sophisticated computational tools, will pave the way to revolutionize the cancer field.

## 7. References

- Alon, U. (2003). Biological networks: the tinkerer as an engineer. *Science*, Vol.301, No.5641, (September 2003), pp. 1866-1867, ISSN 0036-8075
- Anderson, J.C.; Clarke, E.J.; Arkin, A.P. & Voigt, C.A. (2006). Environmentally controlled invasion of cancer cells by engineered bacteria. *Journal of Molecular Biology*, Vol.355, No.4, (January 2006), pp. 619-627, ISSN 00222836
- Andrianantoandro, E.; Basu, S.; Karig, D.K. & Weiss, R. (2006). Synthetic biology: new engineering rules for an emerging discipline. *Molecular Systems Biology* Vol.2, No. 2006.0028, (May 2006), pp. 1-14, ISSN 1744-4292
- Arkin, A.P. (2001). Synthetic cell biology. *Current Opinion in Biotechnology* Vol.12, No.6, (December 2001), pp. 638-644, ISSN 0958-1669
- Bansal, V.; Sakizlis, V.; Ross, R.; Perkins, J.D. & Pistikopoulos, E.N. (2003). New algorithms for mixed-integer dynamic optimization. *Computers and Chemical Engineering* Vol.27, No.5, (May 2003), pp. 647-668, ISSN 0098-1354
- Bar, H.; Yacoby, I. & Benhar, I. (2008). Killing cancer cells by targeted drug-carrying phage nanomedicines. *BMC Biotechnology* Vol.8, No.37, (April 2008), pp. 1-14, ISSN 1472-6750

- Batt, G., Yordanov, B., Weiss, R. & Belta, C. (2007). Robustness analysis and tuning of synthetic gene networks. *Bioinformatics* Vol.23, No.18, (July 2007), pp. 2415-2422, ISSN 1367-4803
- Bedi, D.; Musacchio, T.; Fagbohun, O.A.; Gillespie, J.W.; Deinnocentes, P.; Bird, R.C.; Bookbinder, L.; Torchilin, V.P. & Petrenko, V.A. (2011). Delivery of siRNA into breast cancer cells via phage fusion protein-targeted liposomes. *Nanomedicine: Nanotechnology, Biology, and Medicine*, doi:10.1016/j.nano.2010.10.004, ISSN 1549-9634
- Beisel, C.L. & Smolke, C.D. (2009). Design principles for riboswitch function. *PLoS Computational Biology* Vol.5, No.4, (April 2009), e1000363, pp. 1-14, ISSN 1553-734X
- Benner, S.A. & Sismour, A.M. (2005). Synthetic biology. *Nature Reviews Genetics* Vol.6, No.7, (July 2005), pp. 533-543, ISSN 1471-0056
- Brown, K.C. (2010). Peptidic tumor targeting agents: the road from phage display peptide selections to clinical applications. *Current Pharmaceutical Design* Vol.16, No.9, (March 2010), pp. 1040-1054, ISSN 1381-6128
- Brustad, E.M. & Arnold, F.H. (2010). Optimizing non-natural protein function with directed evolution. *Current Opinion in Chemical Biology* Vol.15, No.2, (April 2010), pp. 1-10, ISSN 1367-5931
- Bulter, T.; Lee, S.G.; Wong, W.W.; Fung, E.; Connor, M.R. & Liao, J.C. (2004). Design of artificial cell-cell communication using gene and metabolic networks. *PNAS* Vol.101, No.8 (February 2004), pp. 2299-2304, ISSN 0027-8424
- Bumpus, S.; Evans, B.; Thomas, P.; Ntai, I. & Kelleher, N. (2009). A proteomics approach to discovering natural products and their biosynthetic pathways. *Nature Biotechnology* Vol.27, No.10, (September 2009), pp. 951-956, ISSN 1087-0156
- Canton, B.; Labno, A. & Endy, D. (2008). Refinement and standardization of synthetic biological parts and devices. *Nature Biotechnology* Vol.26, No. 7, (July 2008), pp. 787-793, ISSN 1087-0156
- Carothers, J.M.; Goler, J.A. & Keasling, J.D. (2009). Chemical synthesis using synthetic biology. *Current Opinion in Biotechnology* Vol.20, No.4, (August 2009), pp. 498-503, ISSN 0958-1669
- Chan, L.Y.; Kosuri, S. & Endy, D. (2005). Refactoring bacteriophage T7. *Molecular Systems Biology* Vol.1, No. 2005.0018, (September 2005), pp. 1-10, ISSN 1744-4292
- Cooling, M.T.; Hunter, P. & Crampin, E.J. (2008). Modelling biological modularity with CellML. *IET Systems Biology* Vol.2, No.2, (March 2008), pp. 73-79, ISSN 1751-8849
- Cramer, A.; Raillard, S.A.; Bermudez, E. & Stemmer, W.P. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* Vol.391, No.6664, (January 1998), pp. 288-291, ISSN 0028-0836
- Czar, M.J.; Cai, Y. & Peccoud, J. (2009). Writing DNA with GenoCAD. *Nucleic Acids Research* Vol.37, No.2, (May 2009), pp. W40-W47, ISSN 0305-1048
- Dasika, M.S. & Maranas, C.D. (2008). OptCircuit: an optimization based method for computational design of genetic circuits. *BMC Systems Biology* Vol.2, No.24, pp. 1-19, ISSN 1752-0509
- Datsenko, K.A. & Wanner, B.L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *PNAS* Vol.97, No. 12, (June 2000), pp. 6640-6645, ISSN 0027-8424



- Dawid, A.; Cayrol, B. & Isambert, H. (2009). RNA synthetic biology inspired from bacteria: construction of transcription attenuators under antisense regulation. *Physical Biology* Vol.6, No. 025007, (July 2009), pp. 1-10, ISSN 1478-3975
- Di Bernardo, D.; Thompson, M.J.; Gardner, T.S.; Chobot, S.E.; Eastwood, E.L.; Wojtovich, A.P.; Elliott, S.E.; Schaus, S.E. & Collins, J.J. (2005). Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. *Nature Biotechnology* Vol.23, No.3, (March 2005), pp. 377-383, ISSN 1087-0156
- Di Ventura, B.; Lemerle, C.; Michalodimitrakis, K. & Serrano, L. (2006). From *in vivo* to *in silico* biology and back. *Nature* Vol.443, No.7111, (October 2006), pp. 527-533, ISSN 0028-0836
- Dougherty, M.J. & Arnold, F.H. (2009). Directed evolution: new parts and optimized function. *Current Opinion in Biotechnology* Vol.20, No.4, (August 2009), pp. 486-491, ISSN 0958-1669
- Du B, Han H, Wang Z, Kuang L, Wang L, Yu L, Wu M, Zhou Z, Qian M. (2010). Targeted drug delivery to hepatocarcinoma *in vivo* by phage-displayed specific binding peptide. *Molecular Cancer Research* Vol.8, No.2, (February 2010), pp.135-144, ISSN 1541-7786
- Ellis, T.; Wang, X. & Collins, J.J. (2009). Diversity-based, model guided construction of synthetic gene networks with predicted functions. *Nature Biotechnology* Vol.27, No.5, (May 2009), pp. 465- 471, ISSN 1087-0156
- Elowitz, M.B. & Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature* Vol.403, No.6767, (January 2000), pp. 335-338, ISSN 0028-0836
- Elowitz, M.B.; Levine, A.J.; Siggia, E.D. & Swain, P.S. (2002). Stochastic gene expression in a single cell. *Science* Vol.297, No.5584, (August 2002), pp. 1183-1186, ISSN 0036-8075
- Endy, D. (2005). Foundations for engineering biology. *Nature* Vol.438, No. 7067, (November 2005), pp. 449-453, ISSN 0028-0836
- Engler, C.; Gruetzner, R.; Kandzia, R. & Marillonnet, S. (2009). Golden gate shuffling: a one pot DNA shuffling method based on type IIs restriction enzymes. *PLoS ONE* Vol.4, No.5, (May 2009), e5553, pp. 1-9, ISSN 1932-6203
- Eriksson, F.; Tsagotis, P.; Lundberg, K.; Parsa, R.; Mangsbo, S.M.; Persson, M.A.; Harris, R.A. & Pisa, P.J. (2009). Tumor-specific bacteriophages induce tumor destruction through activation of tumor-associated macrophages. *The Journal of Immunology* Vol.182, No.5, (March 2009), pp. 3105-3111, ISSN 0022-1767
- Ferreira, C.S.; Matthews, C.S. & Missailidis, S. (2006). DNA aptamers that bind to MUC1 tumour marker: design and characterization of MUC1-binding single-stranded DNA aptamers. *Tumour Biology* Vol.27, No.6, (October 2006), pp. 289-301, ISSN 1010-4283
- Fisher, J. & Henzinger, T.A. (2007). Executable cell biology. *Nature Biotechnology* Vol.25, No.11, (November 2007), pp. 1239-1249, ISSN 1087-0156
- Forbes, N.S. (2010). Engineering the perfect (bacterial) cancer therapy. *Nature Reviews Cancer* Vol.10, No.11, (October 2010), pp. 785-794, ISSN 1474-175X
- Francois, P. & Hakim, V. (2004). Design of genetic networks with specified functions by evolution *in silico*. *PNAS* Vol.101, No.2, (January 2004), pp. 580-585, ISSN 0027-8424



- Funahashi, A.; Morohashi, M. & Kitano, H. (2003). CellDesigner: a process diagram editor for gene-regulatory and biochemical networks. *BIOSILICO* Vol.1, No.5, (November 2003), pp. 159–162, ISSN 1478-5382
- Fung, E.; Wong, W.W.; Suen, J.K.; Bulter, T.; Lee, S.G. & Liao, J.C. (2005). A synthetic gene metabolic oscillator. *Nature* Vol.435, No.7038, (May 2005), pp. 118–122, ISSN 0028-0836
- Gao, C.; Mao, S.; Ronca, F.; Zhuang, S.; Quaranta, V.; Wirsching, P. & Janda, K.D. (2003). De novo identification of tumor-specific internalizing human antibody-receptor pairs by phage-display methods. *Journal of Immunological Methods*. Vol.274, No.1-2, (March 2003), pp. 185-197, ISSN 0022-1759
- Gardner, T.S.; Cantor, C.R. & Collins, J.J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature* Vol.403, No.6767, (January 2000), pp. 339-342, ISSN 0028-0836
- Garst, A.D. & Batey, R.T. (2009). A switch in time: detailing the life of a riboswitch. *Biochimica Biophysica Acta* Vol.1789, No.9-10, (September-October 2009), pp. 584-591, ISSN 0006-3002
- Goler, J.A. (2004). A design and simulation tool for synthetic biological systems. Cambridge, MA: MIT
- Golynskiy, M.V. & Seelig, B. (2010). De novo enzymes: from computational design to mRNA display. *Trends in Biotechnology* Vol.28, No.7, (July 2010), pp. 340-345, ISSN 0167-7799
- Grunberg, R.; Nilges, M. & Leckner, J. (2007). Biskit – a software platform for structural bioinformatics. *Bioinformatics* Vol.23, No.6, (March 2007), pp. 769-770, ISSN 1367-4803
- Ham, T.S.; Lee, S.K.; Keasling, J.D. & Arkin, A.P. (2006). A tightly regulated inducible expression system utilizing the fim inversion recombination switch. *Biotechnology and Bioengineering* Vol.94, No.1, (May 2006), pp. 1–4, ISSN 0006-3592
- Hartley, J.L. (2003). Use of the gateway system for protein expression in multiple hosts. *Current Protocols in Protein Science* Chap5: Unit 5.17
- Hawkins, T.; Luban, S. & Kihara, D. (2006). Enhanced automated function prediction using distantly related sequences and contextual association by PFP. *Protein Science* Vol.15, No.6, (June 2006), pp. 1550-1556, ISSN 1469-896X
- Hianik, T. & Wang, J. (2009). Electrochemical aptasensors – recent achievements and perspectives. *Electroanalysis* Vol.21, No.11, (June 2009), pp. 1223-1235, ISSN 1521-4109
- Hill, A.D.; Tomshine, J.R.; Weeding, E.M.; Sotiropoulos, V. & Kaznessis, Y.N. (2008). SynBioSS: the synthetic biology modeling suite. *Bioinformatics* Vol.24, No.21, (November 2008), pp. 2551-2553, ISSN 1367-4803
- Hofacker, I.L. (2003). Vienna RNA secondary structure server. *Nucleic Acids Research* Vol.31, No.13, (July 2003), pp. 3429-3431, ISSN 0305-1048
- Isaacs, F.J.; Dwyer, D.J. & Collins, J.J. (2006). RNA synthetic biology. *Nature Biotechnology* Vol.24, No.5, (May 2006), pp. 545-554, ISSN 1087-0156
- Isalan, M.; Lemerle, C.; Michalodimitrakis, K.; Horn, C.; Beltrao, P.; Raineri, E.; Garriga-Canut, M.; Serrano, L. (2008). Evolvability and hierarchy in rewired bacterial gene networks. *Nature* Vol.452, No.7189, (April 2008), pp. 840–845, ISSN 0028-0836

- Jemal, A.; Bray, F.; Center, M.M.; Ferlay, J.; Ward, E. & Forman, D. (2011). Global cancer statistics. *CA Cancer Journal for Clinicians* Vol.61, No.2, (March-April 2011), pp. 69-90, ISSN 0007-9235
- Kaiser, J. (2005). Gene therapy. Putting the fingers on gene repair. *Science* Vol.310, No.5756, (December 2005), pp.1894-1896, ISSN 0036-8075
- Kaznessis, Y.N. (2009). Computational methods in synthetic biology. *Biotechnology Journal* Vol.4, No.10, (October 2009), pp.1392-1405, ISSN 1860-7314
- Kelly, J.; Rubin, A.J.; Davis, J. II; Ajo-Franklin, C.M.; Cumbers, J.; Czar, M.J.; de Mora, K.; Gliberman, A.I.; Monie, D.D. & Endy, D. (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *Journal of Biological Engineering* Vol.3, No.4, (March 2009), pp. 1-13, ISSN 1754-1611
- Kirkpatrick, S.; Gelatt, C.D. Jr. & Vecchi, M.P. (1983). Optimization by Simulated Annealing. *Science* Vol.220, No.4598, (May 1983), pp. 671-680, ISSN 0036-8075
- Kitano, H. (2002). Systems biology: a brief overview. *Science* Vol.295, No.5560, (March 2002), pp. 1662-1664, ISSN 0036-8075
- Kizer, L.; Pitera, D.J.; Pfleger, B.F. & Keasling, J.D. (2008). Application of functional genomics to pathway optimization for increased isoprenoid production. *Applied and Environmental Microbiology* Vol.74, No.10, (May 2008), pp. 3229-3241, ISSN 0099-2240
- Kobayashi, H.; Kaern, M.; Araki, M.; Chung, K.; Gardner, T.S.; Cantor, C.R. & Collins, J.J. (2004). Programmable cells: interfacing natural and engineered gene networks. *PNAS* Vol.101, No.22, (June 2004), pp. 8414-8419, ISSN 0027-8424
- Koide, T.; Pang, W.L. & Baliga, N.S. (2009). The role of predictive modeling in rationally reengineering biological systems. *Nature Reviews Microbiology* Vol.7, No.4, (April 2009), pp. 297-305, ISSN 1740-1526
- Kosuri, S.; Kelly, J.R. & Endy, D. (2007). TABASCO: a single molecule, base pair resolved gene expression simulator. *BMC Bioinformatics* Vol.8, No.480, (December 2007), pp. 1-15, ISSN 1471-2105
- Lartigue, C.; Glass, J.I.; Alperovich, N.; Pieper, R.; Parmar, P.P.; Hutchison III, C.A.; Smith, H.O. & Venter, J.C. (2007). Genome transplantation in bacteria: changing one species to another. *Science* Vol.317, No.5838, (August 2007), pp. 632-638, ISSN 0036-8075
- Leonard, E.; Nielsen, D.; Solomon, K. & Prather, K.J. (2008). Engineering microbes with synthetic biology frameworks. *Trends in Biotechnology* Vol.26, No.12, (December 2008), pp. 674-681, ISSN 0167-7799
- Li K, Chen Y, Li S, Nguyen HG, Niu Z, You S, Mello CM, Lu X, Wang Q. (2010). Chemical modification of M13 bacteriophage and its application in cancer cell imaging. *Bioconjugate Chemistry* Vol.21, No.7, (December 2010), pp. 1369-1377, ISSN 1043-1802
- Loettgers, A. (2007). Model organisms and mathematical and synthetic models to explore regulation mechanisms. *Biological Theory* Vol.2, No.2, (December 2007), pp. 134-142, ISSN 1555-5542
- Mandell, J.G. & Barbas, C.F. III (2006). Zinc Finger Tools: custom DNA binding domains for transcription factors and nucleases. *Nucleic Acids Research* Vol.34, (July 2006), pp. W516-523, ISSN 0305-1048

- Marchisio, M.A. & Stelling, J. (2008). Computational design of synthetic gene circuits with composable parts. *Bioinformatics* Vol.24, No.17, (September 2008), pp.1903-1910, ISSN 1367-4803
- Marchisio, M.A. & Stelling, J. (2009). Computational design tools for synthetic biology. *Current Opinion in Biotechnology* Vol.20, No.4, (August 2009), pp. 479-485, ISSN 0958-1669
- Marguet, P.; Balagadde, F.; Tan, C. & You, L. (2007). Biology by design: reduction and synthesis of cellular components and behavior. *Journal of the Royal Society Interface* Vol.4, No.15, (August 2007), pp. 607-623, ISSN 1742-5689
- Markham, N.R. & Zuker, M. (2008). UNAFold: software for nucleic acid folding and hybridization. *Methods Molecular Biology* Vol.453, No.1, pp. 3-31, ISSN 1064-3745
- Martin, V.J.; Pitera, D.J.; Withers, S.T.; Newman, J.D. & Keasling, J.D. (2003). Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nature Biotechnology* Vol.21, No.7, (July 2003), pp.796-802, ISSN 1087-0156
- Matsuoka, Y.; Ghosh, S. & Kitano, H. (2009). Consistent design schematics for biological systems: standardization of representation in biological engineering. *Journal of the Royal Society Interface* Vol.6, No.4, (August 2009), pp. S393-S404, ISSN 1742-5689
- McArthur IV, G.H. & Fong, S.S. (2010). Toward engineering synthetic microbial metabolism. *Journal of Biomedicine and Biotechnology* doi:10.1155/2010/459760, ISSN 1110-7243
- Mirschel, S.; Steinmetz, K.; Rempel, M.; Ginkel, M. & Gilles, E.D. (2009). PROMOT: modular modeling for systems biology. *Bioinformatics* Vol.25, No.5, (March 2009), pp. 687-689, ISSN 1367-4803
- Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S. & Olson, A.J. (2009). AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *Journal of Computational Chemistry* Vol.30, No.16, (December 2009), pp. 2785-2791, ISSN 0192-8651
- Mueller, S.; Coleman, J.R. & Wimmer, E. (2009). Putting synthesis into biology: a viral view of genetic engineering through de novo gene and genome synthesis. *Chemistry & Biology* Vol.16, No.3, (March 2009), pp. 337-347, ISSN 1074-5521
- Neumann, H. & Neumann-Staubitz, P. (2010). Synthetic biology approaches in drug discovery and pharmaceutical biotechnology. *Applied Microbiology and Biotechnology* Vol.87, No.1, (June 2010), pp. 75-86, ISSN 0175-7598
- Nielsen, J. (2001). Metabolic engineering. *Applied Microbiology and Biotechnology* Vol.55, No.3, (April 2001), pp. 263-283, ISSN 0175-7598
- Park, J.H.; Lee, S.Y.; Kim, T.Y. & Kim, H.U. (2008). Application of systems biology for bioprocess development. *Trends in Biotechnology* Vol.26, No.8, (August 2008), pp. 404-412, ISSN 0167-7799
- Pawelek, J.; Low, K. & Bermudes, D. (2003). Bacteria as tumour-targeting vectors. *Lancet Oncology* Vol.4, No.9, (September 2003), pp. 548-556, ISSN 1470-2045
- Pedersen, M. & Phillips, A. (2009). Towards programming languages for genetic engineering of living cells. *Journal of the Royal Society Interface* Vol.6, No.4, (August 2009), pp. S437-S450, ISSN 1742-5689

- Prather, K. & Martin, C.H. (2008). De novo biosynthetic pathways: rational design of microbial chemical factories. *Current Opinion in Biotechnology* Vol.19, No.5, (October 2008), pp. 468–474, ISSN 0958-1669
- Price, N.D. & Shmulevich, I. (2007). Biochemical and statistical network models for systems biology. *Current Opinion in Biotechnology* Vol.18, No.4, (August 2007), pp. 365–370, ISSN 0958-1669
- Purnick, P.E. & Weiss, R. (2009). The second wave of synthetic biology: from modules to systems. *Nature Reviews Molecular Cell Biology* Vol.10, No.6, (June 2009), pp. 410–422, ISSN 1471-0072
- Quan, J. & Tian, J. (2009). Circular polymerase extension cloning of complex 1 gene libraries and pathways. *PLoS ONE* Vol.4, No.7, (July 2009), e6441, pp.1–6, ISSN 1932-6203
- Richardson, S.M.; Wheelan, S.J.; Yarrington, R.M. & Boeke, J.D. (2006). GeneDesign: rapid, automated design of multikilobase synthetic genes. *Genome Research* Vol.16, No.4, (April 2006), pp. 550–556, ISSN 1088-9051
- Ritchie, D.W. (2008). Recent progress and future directions in protein–protein docking. *Current Protein & Peptide Science* Vol.9, No.1, (February 2008), pp. 1–15, ISSN 1389-2037
- Ro, D.K.; Paradise, E.M.; Ouellet, M.; Fisher, K.J.; Newman, K.L.; Ndungu, J.M.; Ho, K.A.; Eachus, R.A.; Ham, T.S.; Kirby, J.; Chang, M.C.; Withers, S.T.; Shiba, Y.; Sarpong, R. & Keasling, J.D. (2006). Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* Vol.440, No.7086, (April 2006), pp. 940–943, ISSN 0028-0836
- Rocha, I.; Forster, J. & Nielsen, J. (2008). Design and application of genome-scale reconstructed metabolic models. *Methods in Molecular Biology* Vol.416, No.IIIB, pp. 409–431, ISSN 1064-3745
- Rocha, I.; Maia, P.; Evangelista, P.; Vilaça, P.; Soares, S.; Pinto, J.P.; Nielsen, J.; Patil, K.R.; Ferreira, E.C. & Rocha, M. (2010). OptFlux: an open-source software platform for in silico metabolic engineering. *BMC Systems Biology* Vol.4, No. 45, pp. 1–12, ISSN 1752-0509
- Rodrigo, G.; Carrera, J. & Jaramillo, A. (2007a). Asmparts: assembly of biological model parts. *Systems and Synthetic Biology* Vol.1, No.4, (December 2007), pp. 167–170, ISSN 1872-5325
- Rodrigo, G.; Carrera, J. & Jaramillo, A. (2007b). Genetdes: automatic design of transcriptional networks. *Bioinformatics* Vol.23, No.14, (July 2007), pp.1857–1858, ISSN 1367-4803
- Rodrigues, L.R.; Teixeira, J.A.; Schmitt, F.; Paulsson, M. & Lindmark Måsson, H. (2007). The role of osteopontin in tumour progression and metastasis in breast cancer. *Cancer Epidemiology Biomarkers & Prevention* Vol.16, No.6, (June 2007), pp. 1087–1097, ISSN 1055-9965
- Saito, H. & Inoue, T. (2009). Synthetic biology with RNA motifs. *The International Journal of Biochemistry & Cell Biology* Vol.41, No.2, (February 2009), pp. 398–404, ISSN 1357-2725
- Salis, H.M., Mirsky, E.A. & Voigt, C.A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. *Nature Biotechnology* Vol.27, No.10, (October 2009), pp. 946–950, ISSN 1087-0156



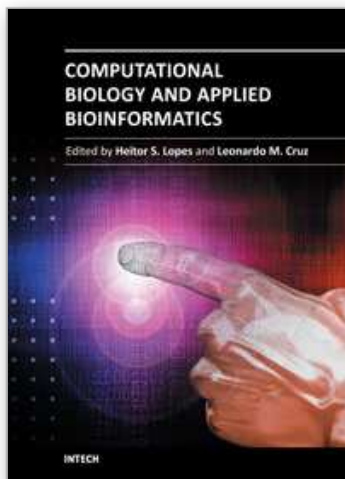
- Segre, D.; Vitkup, D. & Church, G.M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *PNAS* Vol.99, No.23, (November 2001), pp. 15112-15117, ISSN 0027-8424
- Shetty, R.P.; Endy, D. & Knight, T.F. Jr. (2008). Engineering BioBrick vectors from BioBrick parts. *Journal of Biological Engineering* Vol.2, No.1, (April 2008), pp.5-17 ISSN 1754-1611
- Simons, K.T.; Bonneau, R.; Ruczinski, I. & Baker, D. (1999). Ab initio protein structure prediction of CASP III targets using ROSETTA. *Proteins* Vol.3, pp. 171-176, ISSN 0887-3585
- Steinmetz, N.F. (2010). Viral nanoparticles as platforms for next-generation therapeutics and imaging devices. *Nanomedicine: Nanotechnology, Biology, and Medicine* Vol.6, No.5, (October 2010), pp. 634-641, ISSN 1549-9634
- Stoltenburg, R.; Reinemann, C. & Strehlitz, B. (2007). SELEX – A (r)evolutionary method to generate high-affinity nucleic acid ligands. *Biomolecular Engineering* Vol.24, No.4, (October 2007), pp. 381–403, ISSN 1389-0344
- Topp, S. & Gallivan, J.P. (2007). Guiding bacteria with small molecules and RNA. *Journal of the American Chemical Society* Vol.129, No.21, (May 2007), pp. 6807-6811, ISSN 0002-7863
- Tyo, K.E.; Alper, H.S. & Stephanopoulos, G. (2007). Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends in Biotechnology* Vol.25, No.3, (March 2007), pp. 132-137, ISSN 0167-7799
- Tyo, K.E.J.; Ajikumar, P.K. & Stephanopoulos, G. (2009). Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nature Biotechnology* Vol.27, No.8, (August 2009), pp. 760–765, ISSN 1087-0156
- Villalobos, A.; Ness, J.E.; Gustafsson, C.; Minshull, J. & Govindarajan, S. (2006). Gene Designer: a synthetic biology tool for constructing artificial DNA segments. *BMC Bioinformatics* Vol.7, No.285, (June 2006), pp.1-8, ISSN 1471-2105
- Voloshchuk, N. & Montclare, J.K. (2010). Incorporation of unnatural amino acids for synthetic biology. *Molecular Biosystems* Vol.6, No.1, (January 2010), pp. 65-80, ISSN 1742-2051
- Wang, H.H.; Isaacs, F.J.; Carr, P.A.; Sun, Z.Z.; Xu, G.; Forest, C.R. & Church, G.M. (2009). Programming cells by multiplex genome engineering and accelerated evolution. *Nature* Vol.460, No.7257, (August 2009), pp. 894–898, ISSN 0028-0836
- Wierling, C.; Herwig, R. & Lehrach, H. (2007). Resources, standards and tools for systems biology. *Briefings in Functional Genomics and Proteomics* Vol.6, No.3, (September 2007), pp. 240-251, ISSN 1473-9550
- Win, M.N. & Smolke, C.D. (2007). From the cover: a modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *PNAS* Vol.104, No.36, (September 2007), pp. 14283–14288, ISSN 0027-8424
- Winkler, W.C. & Breaker, R.R. (2005). Regulation of bacterial gene expression by riboswitches. *Annual Review of Microbiology* Vol.59, pp. 487-517, ISSN 0066-4227
- Xu, J.; Li, M.; Kim, D. & Xu, Y. (2003). RAPTOR: optimal protein threading by linear programming. *Journal of Bioinformatics and Computational Biology* Vol.1, No.1, (April 2003), pp. 95-117, ISSN 0219-7200



- Yokobayashi, Y.; Weiss, R. & Arnold, F.H. (2002). Directed evolution of a genetic circuit. *PNAS* Vol.99, No.26, (September 2002), pp. 16587-16591, ISSN 0027-8424
- You, L. (2004). Toward computational systems biology. *Cell Biochemistry and Biophysics* Vol.40, No.2, pp. 167-184, ISSN 1085-9195

IntechOpen

IntechOpen



## **Computational Biology and Applied Bioinformatics**

Edited by Prof. Heitor Lopes

ISBN 978-953-307-629-4

Hard cover, 442 pages

**Publisher** InTech

**Published online** 02, September, 2011

**Published in print edition** September, 2011

Nowadays it is difficult to imagine an area of knowledge that can continue developing without the use of computers and informatics. It is not different with biology, that has seen an unpredictable growth in recent decades, with the rise of a new discipline, bioinformatics, bringing together molecular biology, biotechnology and information technology. More recently, the development of high throughput techniques, such as microarray, mass spectrometry and DNA sequencing, has increased the need of computational support to collect, store, retrieve, analyze, and correlate huge data sets of complex information. On the other hand, the growth of the computational power for processing and storage has also increased the necessity for deeper knowledge in the field. The development of bioinformatics has allowed now the emergence of systems biology, the study of the interactions between the components of a biological system, and how these interactions give rise to the function and behavior of a living being. This book presents some theoretical issues, reviews, and a variety of bioinformatics applications. For better understanding, the chapters were grouped in two parts. In Part I, the chapters are more oriented towards literature review and theoretical issues. Part II consists of application-oriented chapters that report case studies in which a specific biological problem is treated with bioinformatics tools.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Lígia R. Rodrigues and Leon D. Kluskens (2011). Synthetic Biology & Bioinformatics Prospects in the Cancer Arena, Computational Biology and Applied Bioinformatics, Prof. Heitor Lopes (Ed.), ISBN: 978-953-307-629-4, InTech, Available from: <http://www.intechopen.com/books/computational-biology-and-applied-bioinformatics/synthetic-biology-bioinformatics-prospects-in-the-cancer-arena>

**INTECH**  
open science | open minds

#### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

#### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen