FEBS Letters xxx (2012) xxx–xxx

journal homepage: www.FEBSLetters.org

Review

Scaling up synthetic biology: Do not forget the chassis

Antoine Danchin *

AMAbiotics SAS, Bâtiment G1, 2 rue Gaston Crémieux, 91057 Evry Cedex, France
Department of Biochemistry, Li KaShing Faculty of Medicine, Hong Kong University, 21 Sassoon Road, Pokfulam, Hong Kong

ARTICLE INFO

Article history:
Received 25 November 2011
Revised 16 December 2011
Accepted 19 December 2011
Available online xxxx

Edited by Thomas Reiss und Wilhelm Just

Keywords:
Genetic program
NanoRNase
Molecular chaperone
Ageing
Logicome
Paleome
Cenome

ABSTRACT

Using comparative genomics and functional analysis, this work summarises how the cell's genome is organised, with emphasis on the importance of the cell's chassis. Some discrete but important engineering constraints are reviewed, beginning with the need for scaffolds, as well as the question posed by the difficult task of putting a very long random thread (DNA) into a limited volume. Subsequently, to illustrate overlooked essential functions, we show the importance of safety valves, as well as the need to cope with leftovers. The third section discusses how transplantation experiments point out a remarkable feature of the cell factory: the program replicates (makes *identical* copies of itself), whereas the cell reproduces (makes *similar* copies of itself), placing in the limelight the role of informational maintenance. A final section identifies the need to put together a globally linear behaviour of the cell with intrisically non-linear genetic constructs. The discussion ends with the central question of evolvability of artificial constructs and to suggest that combining in vivo Synthetic Biology with biochemical reactors might be an efficient way forward.

1. Introduction

The present avatar of «Synthetic Biology» (SB) assumes that we know enough of what life is to allow us to construct life from scratch, or, at least, to modify existing cells and organisms so that they work as cell factories. With this view SB puts together two separate entities, a program (the conceptual extension of the genetic program) and a chassis (the conceptual extension of the living cell). Yet a rapid browsing of publications in the domain indicates that the vast majority of SB-related work deals with the program, not the chassis. Most investigators do as if the recipient cell would be nice enough to accommodate entirely artificial constructs and behave as expected, producing the right products, with the right yield, at the right time. Several hard constraints however make this dream difficult, if not impossible to achieve. We review here some of the specific features of the cell chassis as well as hidden constraints that must be understood before we really go for large-scale industrial success stories.

The originality of contemporary SB (for the history of the concept see [1–3], and in French, an account of Stéphane Leduc's work [4]) is that it places engineering at its heart. In concrete terms, what would we need if we were to construct an organism that would carry over a synthetic genome and stay alive for a long time? Would scaling up synthetic processes be possible, and to what extent [5]? It requires us to try and make a thorough inventory of functions, taking care not to forget unobtrusive but essential ones, to see whether we do not miss important points. To this aim, we analyse the two independent components [6] of the constructs: the chassis (the cell) and its program and see how they ought to come together smoothly.

The present review is split into four parts. Using comparative genomics and functional analysis [7], we first summarise how the cell's genome is organised, with emphasis on the cell as a factory. We then illustrate some unobtrusive but important engineering constraints that need to be considered. To this aim, we begin with the need for scaffolds, as well as the hurdle created by the difficult task of putting a very long random thread into a limited volume. Then, as an illustration of overlooked essential functions, we show the importance of safety valves, as well as the need to cope with the leftovers that are inevitably created during maintenance processes. The third section discusses how genome transplantation experiments point out a remarkable feature of the cell factory: the program replicates (makes *identical* copies of itself), whereas the cell reproduces (makes *similar* copies of itself) [8]. This subtle difference in processes that are usually mistaken for a single one introduces the category «information» as a currency of reality essential to understand how life develops. The fourth section identifies the need to put together a globally linear behaviour of the cell's physiology with intrinsically non-linear genetic constructs. This allows us to end our discussion with the central question of evolvability of artificial constructs and to suggest that combining

* Fax: +331 6087 5303.
E-mail addresses: antoine.danchin@normalesup.org, adanchin@hkucc.hku.hk

0014-5793/$36.00 © 2011 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.febslet.2011.12.024

in vivo SB with biochemical reactors might be an efficient engineering way forward.

2. Functional analysis of the genome

More than half a century of molecular genetic studies have unraveled the bulk of the functions required to make a living cell. The belief that we already know everything about life is at the root of the present success of the SB stance, which implies that we are in a position to reconstruct life, or even construct novel forms of life using our previous abstract knowledge, that can be summarised in the way we consider the organisation of bacterial genomes.

2.1. The minimal genome

At the onset of genome projects, back in the mid-1980s, the idea already prevailed that we knew the vast majority of the functions required to sustain life [9]. In particular it was possible to establish a list of the minimum number of functions that would be necessary, if perhaps not sufficient, to account for the properties of living systems. This list was published in the white paper under the helm of André Goffeau meant to justify the funding of genome projects by the European Commission. The minimal genome then comprised approximately 400–500 kb of DNA, with a core machinery (replication–transcription–translation: the core of the molecular biology «dogma») made of about 250 genes [10] (Table 1). Remarkably, sequencing of the shortest genome of an autonomous bacterium later substantiated this figure, with many of its 524 (RNA and protein-coding) genes coding for known functions (but still with a significant number of unknowns) [11].

Identifying essential functions, though, does not tell us how they are implemented: it is capital to remember that while specific functions may be omnipresent, the corresponding structures may (and do) differ from one organism to another one [12]. Furthermore, when tested for dispensability, genes and functions have been assayed over a small number of generations and under laboratory (i.e. stable) conditions. Many more functions than the known ones might need to be taken into account to account for sustained life in the long term (or across many generations).

In any event, the minimal set thus established had a limitation. It assumed that all the building blocks needed to construct the nuts and bolts (metabolites) that make the cell factory are provided by the environment. Many more functions are required if the environment is purely mineral. In running a cell, some 800–1500 metabolites are used, sometimes only as intermediates. This corresponds to some 800–1500 reactions that must run by specific enzymes. As a consequence the minimal genome for a cell exploiting a purely mineral environment would require up to 2000 genes. Indeed, chemolithoautotrophic bacteria such as *Thiobacillus* sp. have a genome of 2.9 Mb [13]. Some cyanobacteria, using carbon dioxide as carbon source, such as *Prochlorococcus marinus* strains have a genome in the 1.6–2.7 Mb range [14].

Table 1
The minimal set of genes postulated in 1989 to justify financing genome projects by the European Commission [10]. The average length of the proteins involved was assumed to be approximately 300 residues (coded by a 1 kb long gene).

Process	Structure	Length (kb)
Replication	DNA wielding	40
Transcription	Transcription + coupling with translation	30
Translation	Ribosome: ribosomal RNA + 50–60 ribosomal proteins	60
	tRNAs + tRNA loading + polypeptide synthesis	80
Core metabolism	Building blocks and coenzymes	200
Transport	Import and export	
Energy management	ATP synthesis and electron transfers	
Specific casings	Creation of an envelope	100

2.2. The paleome and the cenome

Yet, the genome of bacteria that do not rely on complex media is often much larger. For example *Cupriavidus* (*Ralstonia*) *metalliduran*s (formerly *Alacaligenes eutrophus*), a facultative chemiolithoautotroph proposed as a source for enzymes using hydrogen in a fuel cell [15], has a long genome. It is approximately 7000 kb long, made of two chromosomes and two megaplasmids [16], that is, considerably longer than the minimal size just discussed. As a matter of fact, the majority of known free-living bacteria have genomes of a size in the 3000 kb range, usually as a single chromosome. It is therefore important to analyse the distribution of genes in these genomes, asking first the question, what is an essential gene? A thorough exploration of bacterial genomes showed that, contrary to expectation, no gene is common to all bacteria [12]. As a consequence, it is not possible to identify essential functions by overlapping genomes sequences while looking for orthology. An alternative way is to look for *persistent* genes, i.e. orthologous genes that tend to be present not in all but in a quorum of genomes [17]. This way, it is possible to identify approximately 500 genes that could be considered as essential not only to sustain life for a few generations, but to sustain life over a long time in a rich environment where the necessary metabolites are supplied [17]. This number increases to some 1500–1800 genes if one takes into account only bacteria that can grow on a limited number of carbon and nitrogen substrates. The corresponding genes tend to be grouped into a small number of clusters in the genome [18].

To account for the remaining genes (usually spanning at least half of the genome's length) we have proposed to split genomes into two parts, the *paleome* and the *cenome* [19]. For a given species, the core persistent functions are coded by genes that make the paleome of the species (this name was chosen because persistent genes code for functions that must have been present since the early times of life [20]). The paleome common to most bacterial clades (approximately 500 genes) comprises genes required for DNA replication, genes for transcription and translation, genes essential for the cell's maintenance, genes for synthesis of an envelope and transport, and a small set of core metabolic genes. The replication process entails genes for DNA replicase, and accessory proteins: DNA ligase, DNA polymerase I, chromosome compaction, and genes allowing chromosome segregation in daughter cells. This makes some 50 genes. The transcription and translation machineries, with coupling factors (Rho and Nus proteins [21]) make the bulk of the core paleome, with approximately 200 genes (including RNA modification genes, as well as a variety of scaffold protein genes, as discussed below). The cell's maintenance is driven by complex machineries that use energy to discriminate between functional and crippled macromolecules and cope with leftovers (see also below). This process is not entirely deciphered, but it is likely to require some 50 genes. Synthesis of an envelope and the division machinery will ask for some 30–50 genes also. A minimum of twenty specific transporters (importers and exporters) will drive exchanges with the environment. Finally, core metabolic processes, including energy production and management (in general respiration and core intermediary metabolism) will ask for 50–100 genes. This list is not very different from that guessed in 1989 (Table 1). If we consider the paleome of bacteria living in a poor environment, most basic metabolites (building blocks and coenzymes) will have to be synthesized de novo, and this will require some 1000 further genes, mainly coding for enzymes of metabolic pathways. It can be expected that in a fairly near future all functions of the paleome will be identified.

By contrast, the cenome complement of a genome is extremely variable and differs from strain to strain in a given species. It is made of the genes that allow the cell to occupy its specific ecological niche. These genes are acquired by horizontal gene transfer from the large pool of genes that is continuously discovered by metagenomic studies. The processes involved (transformation, transduction, integration of prophages and conjugation) make that these genes are coming in genomes as gene clusters. This very large pool of genes, the cenome (as in biocenosis [22]) is sampled as specific subsets in the different strains of the same species, providing the strain with it original properties.

The *pan-genome*, which labels the overlapping genomes of all strains of a given species [23], is the sum of the paleome and of the cenome (i.e. the collection of the paleome complements in each individual strain belonging to that species). Symmetrically, the cenome of a given species is a subset of the corresponding pan-genome, comprising all the genes allowing any strain of that particular species to live in its favoured niche. As an illustration, in a species such as *Escherichia coli* the pan-genome is the union of the genes forming the *E. coli* paleome (approximately 1800 genes [24]) and the cenome of each individual strain. It already comprises some 20 000 genes, with not much levelling off as new strain genomes are sequenced. A particular cenome subset is responsible of the specific phenotype of each particular strain (e.g. see strain O104:H4 that triggered a dangerous outbreak in Germany in may 2011 [25]).

With this view the minimal genome comprises genes coding for the functions of the paleome common to all free-living bacteria. The paleome can be further split into three functional families, replication, constructing biomass, and safety, maintenance and repair.

2.3. Is there a link between the organisation of the genome and the chassis?

This organisation of the genome does not tell much about the way it could be connected to the chassis. Back in 1980 King et al. reviewed the situation created by the gap between knowledge of the sequence and knowledge of the structure of proteins: «*Historically a gap has existed between the study of the one-dimensional organization of hereditary information in genes, and of the three-dimensional organization of macromolecules in biological structures*» [26]. The gap in our knowledge is even deeper when we consider the cell or the organism as a whole. The image of the computer allows us to better visualise the situation. John von Neumann designed a concrete way to construct functioning Turing Machines, today's computers. Central to his approach was the concept of operating system (OS), the core algorithm that was to manage the interaction between the machine and its environment (including its own «organs», such as the keyboard, screen etc) [27]. An OS is a software that allows the computer hardware to communicate and operate the softwares that are linked to the computer, either as internal software or as external devices or programs. As a consequence, to be working an OS needs to know where to find the routines required for interacting with the components of the hardware. This asks for implementation somewhere in its design an in-built image of the machine. Do we see such an image in the way genomes are organised? Analysis of the *mur-fts* clusters in a variety of bacterial genomes suggests that this is indeed the case [19,28]. In complex organisms the situation is even clearer: the body plan is directly related to the way genes are displayed in genomes, not only in position, but in length (the introns play there an important role) [29]. We will come back to this remarkable feature when discussing the relationship between the non-linear logic of gene expression in relation to the linear logic of growth.

3. Specific engineering requirements for the chassis

The functional view explored up to this point combines the outcome of a data-driven approach with the deep reflection of von Neumann. With the steady progresses of classical genetics, functions have been identified after selection of mutants in specific environments, looking for a particular phenotype, or collected at random after heavy mutagenesis and careful observation of the cell's phenotype (for bacteria most generally the phenotype of a colony on a Petri dish). In the mid-1980s the number of mutants collected this way, using mainly model bacteria and the yeast *Saccharomyces cerevisiae*, resulted in such a large family of functions that many authors thought that we already had got the bulk of genes and functions necessary for life. Briefly, if one isolated a mutant and the corresponding gene from any organism, then sequenced the gene – this was tedious at the time but feasible – there was a very high probability that the gene coded for a protein similar to one already identified and deposited in the first data libraries, the EMBL databank and GenBank [9]. This was indeed a strong argument used by many to fight against genome sequencing programmes, thought to be both extremely expensive and useless [30]. In 1991 however, at the meeting *Genome Analysis in the European Community* at Elounda in Crete, the presentation of the complete sequencing of yeast's chromosome III, in parallel with that of a 100 kb continuous segment of the *Bacillus subtilis* genome revealed a completely unexpected observation: more than half of the putative genes carried by these long DNA sequences did not look like anything known. This was the first discovery of genomics, and this opened a question which is still valid today: what are the functions that we failed to predict? Beside entirely novel metabolic pathways, that keep being discovered, some are described in what follows.

3.1. Scaffolding

Within the SB paradigm, the cell factory can be seen as a computer where the program is physically distinct from the machine. But this computer would have a feature not (yet) displayed by man-made computers, it would be able to create a viable offspring [19]. This view of the cell tells us something about the organisation and evolution of its genome. Yet, it does not tell us much about the way the cell manages engineering constraints. The cell's chassis combines management of compartmentalisation (the cell's envelope, intracellular compartments, appendages, but also nanomachines such as the ribosome, ATP synthase, the proteasome, the degradosome and many others [31]) and metabolism (fluxes of molecules for the building up, storage, salvage of the cell's building blocks, catalytic centres and energy management). Nutrients are imported and waste products are exported. However an engineer would see further needs, such as lubrication or safety devices. Thinking in this way is a means to uncover many unexpected functions.

Remarkably, despite the onset of SB the engineering reasoning has still not gained much ground today, when thinking about the chassis. A first functional category that fills the gap comes to mind. When we construct buildings, it is oftentimes not possible to complete the building from bottom up without using a complementary, usually temporary, device, a scaffold. This structure is implemented in the construction of bacteriophages, for example *B. subtilis* phage SP3, where a scaffold stays in the final construct [32]. In bacteriophage lambda, a temporary scaffold allows construction of the phage head. It is subsequently removed, exactly as is the fate of scaffolds in human constructs [33]. This general engineering requirement has been detailed for phage T4 morphogenesis. The process is a model for construction of the complex capsid of

Q1 Please cite this article in press as: Danchin, A. Scaling up synthetic biology: Do not forget the chassis. FEBS Lett. (2012), doi:10.1016/j.febslet.2011.12.024

bacteriophages, with structural and accessory proteins required for proper assembly including a vernier for measuring the tail length [34].

A related discovery followed the discovery of the GroE system in the control of bacteriophage lambda morphogenesis [35]. In a seminal paper Ellis proposed in 1987 the concept of *molecular chaperones*, proteins that would help shaping proteins during their synthesis or subsequently to reach their final functional form [36]. Since then a large number of previously identified proteins with unknown function, but often characterised by the phenotype of mutants (often defect in phage propagation) was uncovered. Novel members of this family keep being discovered (see for review [37,38]). Many are functionally associated to physical stresses, such as heat or cold shock and are involved in the general quality control of the cell's components [39].

3.2. Packaging the genetic program

Computers are not purely abstract entities. They are very concrete. They run programs, and a program needs a physical support. It can be stored on a CD. In real life a CD is deformable, by heat for example. When deformed, and despite the fact that the program it carries remains unaltered, the laser beam meant to read it will not be able to do so. The program will not be usable by the computer. This does not alter the very existence of either the computer or the abstract laws establishing what a computer is (a Turing Machine). This tells us however that in any concrete implementation of the Turing Machine, one cannot completely separate between the hardware and the software [19]. In cells, the support of the genetic program is a DNA molecule, that, as the CD for the computer program, has specific physico-chemical properties. This observation points to an important constraint that may explain why DNA transplantation experiments have not yet spread worldwide.

Cells have been transformed by DNA probably since the outset of life. Scientists have used a variety of methods to use this property: conjugation, transduction, transformation, and, recently, transplantation of a whole synthetic genome [40]. At the time of this review this latter feat has been accomplished only in one laboratory, showing that the technology has still to be considerably improved, were it to reach the level of industrial applications. Indeed, DNA when within cells is under a condensed shape that allows a small volume to accommodate a very long thread. This involves supercoiling, stabilised by a variety of DNA-binding proteins (e.g. H-NS, HU, FIS, CRP and all transcription regulators) as well as active topoisomerases. When condensed the DNA molecule is extremely fragile as a single nick would immediately decondense it, expanding ten times the radius of the volume accomodating the molecule [41]. Even using very mild extraction techniques it is unlikely that DNA purification from cells could prevent formation of such nicks. This would however preclude transplantation into a cell of similar size as that accomodating a condensed chromosome of the same length.

A way out would be to use large cells, or cell's syncytia (cells fused together in a single one). Looking carefully into the technology used to transplant *Mycoplasma capricolum*, it can be seen that polyethylene glycol (PEG) is an ingredient of the transplantation medium. PEG is known to allow cytoplasmic membranes fusion [42]. It is therefore likely that the success of the approach entailed fusion of several cells where decondensed donor DNA could be accommodated. Upon entry into the syncytium the donor DNA is repaired, packaged and expressed. It codes for a restriction nuclease that destroys the host chromosome, and the syncytium is rapidly budding into a progeny, with the replicas of the now condensed donor genome (Fig. 1). Preparing synthetic DNA with the proper shape is therefore essential for future progresses (see for example [43]). Beside chemical techniques promoting transplantation, in vivo approaches can be used. Viruses are cases in point, and they are used for transduction or transformation, e.g. with cosmid libraries. Their very existence requires that they interact with the cell's envelope, and inject their genetic material into the cell. Some viruses can harbour DNA fragments of considerable length, often originating from their hosts: bacteriophage T4 carries many genes matching those of the cell's paleome, including tRNAs, and there exists giant viruses with even more «cellular» genes, which could transduce huge DNA sequences [43]. It is not clear however that they could carry over a whole cell's genome, as the extant ones are always restricted in such a way that they cannot behave as endosymbiotic cells (even as regressed as mitochondria), but only as self-propagating, cell-dependent, viruses.

3.3. Safety valves

Many further essential yet overlooked functions must be taken into account when aiming at constructing a viable cell. It is generally accepted, without analysis of the inevitable consequences in terms of engineering, that cells have both highly specific and active transport systems. This should have triggered a question: what would happen if the cell had constitutive or previously induced transporters suddenly placed in a medium rich in the transported metabolite? The obvious consequence is that the metabolite would rapidly flow in, and frequently lead to an unsustainable osmotic pressure. For an engineer, this would mean that the cell requires specific processes to prevent disruption of the cell's envelope. One way would be rapid polymerisation of the metabolite, storing it while protecting the cell against pressure-driven explosion. Synthesis of glycogen, in the case of glucose, nicely plays this role. But there is no analog of carbohydrate polymers for each type of sugar transported in the cell. Furthermore storage implies a complex organisation within the cell's compartments. Another escape route must be implemented.

An engineer would propose a safety valve: once reaching a threshold level the valve opens and excess metabolite is expelled out of the cell. This requires specific exporters. And indeed cells contain usually a considerable number of so-called «multidrug resistance» permeases often with no known specificity but recruited to export toxic compounds out of the cell (see e.g. [44] for the current view about their role and origin). Within the frame of the present reflection their presence is not a surprise, it is a clear engineering requirement. However, a further subtle requirement must also be implemented: releasing the metabolite in the environment in the form it has entered the cell will create a futile cycle. It is therefore likely that the metabolite is modified before being released outside the cell. This is exactly what happens in the case of lactose in *E. coli*, and this provides a function for lactose transacetylase (LacA), *including the fact that the enzyme has a poor* K_M, a necessity for the cell to avoid exporting the metabolite it needs to build up its biomass, unless it is overflowing in [45]. It is interesting to note that LacA interferes with SB constructs using lactose analogs as inducers [46]. A noteworthy engineering point here is that assuming that enzymes must always be efficient can lead to considerable misunderstanding of the cell's metabolism. This also explains why drug resistance is so ubiquitous and so easy to acquire: most if not all cells have pre-engineered an in-built process to get rid of undesirable metabolites combining metabolite modification (often acetylation) and export.

3.4. Coping with leftovers

A final example of an engineering constraint is the fate of the ubiquitous degradation leftovers. Proteins and RNA age and are degraded when they can no longer be repaired. In particular messenger RNA molecules must turnover, often fairly fast, to allow the cell

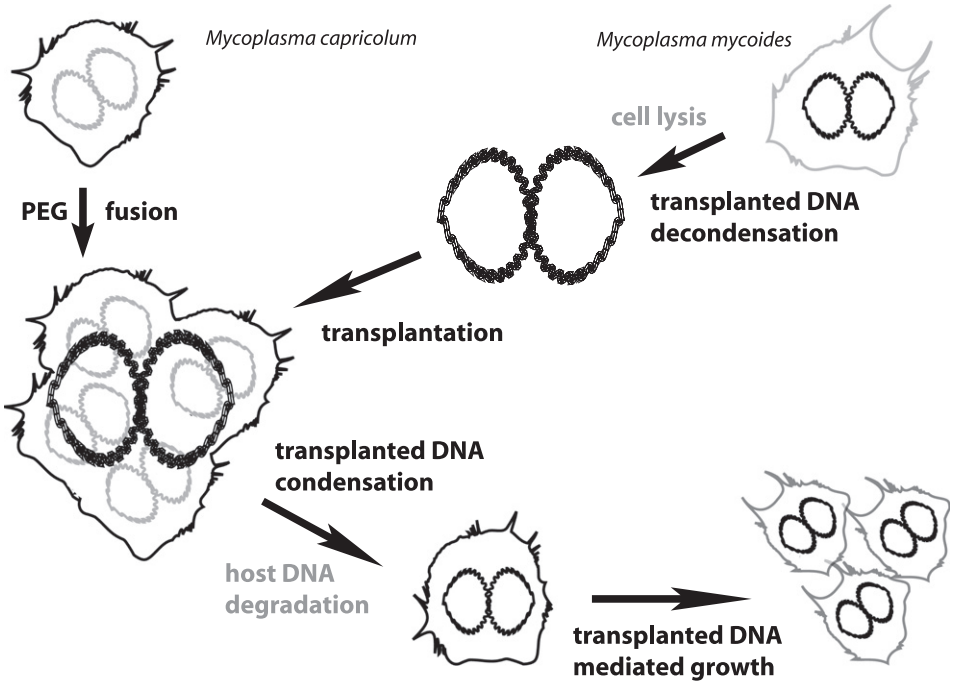

Fig. 1. Chromosome transplantation experiment. To behave as a computer (a Turing Machine) a prerequisite is that the program is physically distinct from the machine that reads it. Lartigue and co-workers [6] have shown that it is possible to transplant the chromosome of *Mycoplasma mycoides* into *M. capricolum*. Yet, because of the extreme fragility of the DNA molecule, it is decondensed as it is damaged by a few nicks, resulting in a molecule that is coiled into a volume much larger than that of the receiving host cell. Using polyethylene-glycol allowed the authors to create a large syncytium that can accommodate the decondensed molecule. This molecule is then repaired and condensed back to its original volume, triggering gene expression, including that of a restriction endonuclease that destroys the host genome. This experiments further illustrates the difference between replication (the transplanted DNA sequence is the same at the outset and at the end of the experiment) and reproduction (the host cell's components differ at the end of the experiment from those of the initial syncytium).

to respond rapidly to variations in its environment. Ribonucleases are either cleaving the mRNA within the molecule, or starting from an end and proceeding via a step by step removal of one nucleotide at a time. During this process the RNA molecule is bound to the enzyme via its charged surface. Once the enzyme nears the end of the molecule these interactions become progressively weaker, and short oligonucleotides (nanoRNAs) detach. They are not innocuous, as their size is such that they will easily enter the transcription «bubble» and jumble the process. A ubiquitous engineering solution exists: specific nanoRNases are present in all organisms, belonging to different descents, but with the same essential function [47]. It is likely that similar processes exist to cope with protein leftovers. Not much work has been devoted to them.

4. Reproduction and replication: cells as inventors

These examples set the stage: many engineering requirements must be implemented to smoothly run the chassis. Beside these functions we also must recognise a specific property of living cells, that differentiates them from standard machines: they make a *young* progeny, and being young implies a noteworthy difference between the parent and its offspring. The program is kept the same but the chassis differs in the parent and its offspring.

4.1. Information of the program, information of the chassis

The transplantation experiment illustrates this remarkable difference. The transplanted program (from *Mycoplasma mycoides*) will initially use the chassis of *M. capricolum* (clearly different from that of *M. mycoides* in biochemical terms). As the cell multiplies, new parts (coded by the transplanted program) will progressively replace the old ones. After some time all the old components have been replaced by similar but not identical components (Fig. 1). This is the living counterpart of the ship that brought Theseus from Crete back to Athens (the Delphic Boat). To be kept as a memorial it was maintained as a ship ready to sail, with new parts replacing the worn out parts, until none of the original ones remained. This process of reproduction was much discussed by ancient Greeks [48]. It illustrates the difference between *replication* (making an exact copy) and *reproduction* (making a similar copy). The program has been replicated, the host chassis has been reproduced. During reproduction something has been conserved, an information, which is beyond the matter of the chassis.

SB implements a program in a chassis. While it is straightforward to exploit the information of the program (this has been done in a large number of studies, generally using Shannon's information, see for review [49]), there is, as yet, no theory that could explore and use the information of the chassis. Yet that the chassis is information-rich is obvious: for the Delphic Boat, for example, oak would be less prone to rot than pine because it is harder. The information «rotting-propensity» has been extracted from a fairly noisy environment in the course of the process of constructing, then maintaining the boat. Note that this may look as if a «poor» information had been extracted in the process, but it could well be that the boat is in an environment where bugs eat oak, and are killed by the resin in pine: that would be clearly positive. This latter type of information, therefore, has to be placed using some kind of measure of distance between the entity considered, and the entities present in its environment. This is particularly important for a cell's progeny, during scaling-up processes, when cells multiply over a large number of generations in an environment that they modify by the change induced by their very presence.

Future work on a novel theory of information needs to be developed in this direction, and epigenetic heredity is perhaps the first place we should investigate in this light. In general, exactly as illustrated by the subtle difference between reproduction and replica-

Q1 Please cite this article in press as: Danchin, A. Scaling up synthetic biology: Do not forget the chassis. FEBS Lett. (2012), doi:10.1016/j.febslet.2011.12.024

tion, there is a systematic mixing up between genetic heredity and epigenetic heredity. The former corresponds to DNA replication; the latter corresponds to reproduction of a particular expression state of DNA. Epigenetic inheritance is much less stable than genetic inheritance, although it can be stable over many generations [50]. Typically cells of the different tissues of a multicellular organism will display a specific expression pattern that is the hallmark of the cells of that particular tissue (skin, liver, blood, etc.). But this can sometimes be reversed, and the general work on dedifferentiation of cells aims at resetting the expression of a given cell type to that of another cell type. This illustrates that there is constant exchange between information of the program and information of the chassis and shows that, in terms of analysis of information, one cannot limit research work to information of the program, as is the present situation.

4.2. Maxwell's demon

How does the cell manage this exchange? Let go back to the paradigm of molecular biology, the idea that molecules represent the lowest significant level needed to account for what life is. James Clerk Maxwell, in his *Theory of Heat* introduced the idea of the "molecular theory of matter", where movement and information are central [51]. Temperature here measures the degree of agitation of a gas: fast when hot, slow when cold. In a gas, if one starts with an asymmetrical distribution, with hot gas molecules in one compartment, and cold gas molecules in a contiguous one, the system will evolve so that the temperature is averaged after some time elapsed with parallel increase in entropy. To create a link between information and entropy, Maxwell explored the idea of a hypothetical being (later named a 'demon') that uses an in-built information-processing ability to reduce the entropy of a homogeneous gas (at a given temperature). Briefly, the demon is able to measure the speed of gas molecules and open or close a door between two compartments as a function of the molecules' speed, keeping them on one side if fast, and on the other side if slow. The demon manipulates an information. Remarkably, his action will build up two compartments, one hot, and one cold, reversing time, and acting apparently against the second principle of thermophysics. How is this possible? Where is the flaw in the reasoning?

Information is also central to computation. The role of thermodynamics in computation has been examined repeatedly over the past half century. The physics of information-processing proposed a considerable variety of attempts to understand how Maxwell's demon could function. A major contribution to this work was the account provided in april 1913 by Marian Smoluchowski, professor at the Jagiellone university in Krakòw. At a lecture in Göttingen attended by the most creative physicists and mathematicians of the time, Smoluchowski gave details of the way Maxwell's demon could be implemented as a trap door, permitting information to be coupled to availability of energy and material states of molecules in the environment [52]. Later on, Szilard proposed in a loose way to account for the relationship between information and entropy [53], and von Neumann in the 1950s followed suit, stating that each logical operation performed in a computer at temperature T must use an energy of kTln2, thereby increasing entropy by kln2 (see [27]). This remained the accepted intuition until the IBM company, which was concerned by the limits this would impose on computation, asked its engineers to explore the situation and possibly propose remedies.

Fortunately for computer sciences (you could not work on the machine you are using if this had reflected reality), and as we shall see for biology, Szilard's intuition proved to be wrong. In the late 50s, construction of computers began to reach an industrial level. Among the many engineering tasks was that of making computers that would be much smaller than the gigantic machines constructed at the time, and also faster machines. If the idea that a logical operation was consuming energy it was obvious that a limit in the space and speed of computation would be rapidly met. A central question was therefore to calculate what would be that limit. In a surprising twist of history of engineering, that was also a revolution in the understanding of the physical world, Rolf Landauer demonstrated in 1961 at IBM that, contrary to intuition, computation could be made to be reversible, hence not to consume any energy [54].

To understand the meaning of this statement, let us summarise the bases of all computations. Three core boolean operations, AND, NOT and REPLICATE are enough to permit all kinds of logical operations. Note that in standard boolean logic REPLICATE is not used, but OR is used instead, with different outcomes in terms of general consequences of what can be performed by the operators. Used in some hardware languages such as Verilog, or the Ada software language, REPLICATE is particularly interesting for allowing fairly straightforward biological intuitions. The operation AND is boolean intersection (multiplication), as we learnt in our first years at school: it takes two binary inputs X and Y and returns the output 1 if and only if both X and Y are 1; otherwise it returns the output 0. Similarly, NOT takes a single binary input X and returns the output 1 if X = 0 and 0 if X = 1. REPLICATE takes a single binary input X and returns two binary outputs, each equal to X. Any boolean function can be constructed by repeated combination of AND, NOT and REPLICATE. Another operation, that can be derived from those, ERASE, is essential to our topic. ERASE is a one-bit logical operation that takes a bit, 0 or 1, and restores it to 0. Concretely, these operations are implemented as 'logic gates'. A logic gate is a physical device that performs a logical operation. Microprocessors are combining millions and even billions of logic gates to perform the complex logical operations that you find in computers such as the one you are using to read this text. SB uses a parallel logical analysis to implement its 'logicome' [55].

In his work, Landauer showed that reversible, one-to-one, logical operations such as NOT can be performed *without consuming energy*. He also showed that irreversible, many-to-one operations such as ERASE require consuming at least kTln2 of energy for each bit of information lost. The core of the argument behind Landauer's theorem can be readily understood. Briefly, when a bit is erased, the information it contains must go somewhere. It has only two possible ways: either it moves to a place in the computer (or of the cell, if we consider cells as computers) corresponding to an observable degree of freedom, such as another place with a known bit in its memory. If so, it has obviously not been erased but merely moved. Or it goes into places with unobservable degrees of freedom such as the microscopic motion of molecules, and this results in an increase of entropy of at least kln2.

In 1973, Bennett extended Landauer's theorem, showing that all computations could be performed using only reversible logical operations, that is, without consuming energy. Briefly, no energy was necessary to create novel information. Only one problem remained: memory had to be erased to allow further computation and this required energy [56]. Where does the energy come from? To perform a logical operation, it is commonly extracted from a store of free energy, then used in the processor that performs the operation, and finally returned to the initial store once the operation has been performed. No energy, in the end, is used in the process. However, to restore the system in its original form, energy must be used. We note here that in usual computers the store is a battery or an outside electric supply, whereas in cells energy is distributed throughout the matter of the cell. This has considerable consequences in the management of information by cells. The property of reversibility has been implemented in real computers under the term "adiabatic logic", and real circuits have been described in details to explain how this works [57].

Q1 Please cite this article in press as: Danchin, A. Scaling up synthetic biology: Do not forget the chassis. FEBS Lett. (2012), doi:10.1016/j.febslet.2011.12.024

4.3. Energy-dependent degradation

The shift from a continuous description of matter to a discontinuous, atomist view, was later on extended to biology with the birth of molecular biology. Note that this took about one century to be accepted. The present situation, where "information" is slowly gaining ground as an authentic currency of biological reality, repeats a similar slow path. Back to Maxwell's demon: In a real computation, errors occur, and to get rid of errors will require an irreversible operation, erasure of the wrong information and replacement by the correct one. Hence, this will result in consuming energy in order to restore the errorless situation. If energy were not consumed, then the system would be able to go backwards in time, and we would have created the perpetual movement. How does this work in reality? The situation is similar to that proposed to be the action of Maxwell's demon: measure, store an information, use it via replication of the measurement to re-establish the initial state, and then erase the memory, to reset the initial state of the demon. Central to this action are two logical processes, REPLICATE and ERASE. Among biological functions that would fulfill these requirements we need therefore to look for degradation systems that consume energy. We shall dwell on this point later on.

If the error rate is x bits per second, for example, then error-correcting processes can be used to detect those errors and reject them to the environment at an energy cost of x kT ln2J s–1, where T is the temperature of the environment. In fact, biological processes, even at the microscopic level, do not proceed bit by bit, but, rather are highly redundant and simultaneously change a fairly large number of bits. This is because at 300 K, the average temperature of life environment, the thermal noise is fairly large so that redundancy is necessary to increase the signal to noise ratio. And the usual "quantum" of energy used is that of hydrolysis of a "energy-rich" phosphate bond, typically hydrolysis of ATP to ADP or GTP to GDP.

While these types of processes have remained conceptual and have not yet been presented as *concrete* illustrations of Maxwell's demon, we have in biology a wealth of examples illustrating behaviours of that type. As already stated, many cell's components have to be continuously replaced. Non-functional proteins can be either repaired [58–60] or refolded, and this costs energy. When neither is possible they are degraded. Degradation enzymes, often hydrolytic enzymes, are essential at this stage. Hydrolysis is exothermic (therefore should produce energy), yet many degradation enzymes *consume energy* – an apparent waste. This reminds us of the remark made by Hopfield, that in order to identify important unexpected functions, we should explore reactions that use energy in an apparently expletive way: "*known reactions which otherwise appear to be useless or deleterious complications*" [61]. And, indeed, in the protein translation process, a proofreading step, using protein EFTu bound to charged transfer RNA, tests whether the tRNA can read correctly the codon immediately available after the tRNA carrying the growing polypeptide, and hydrolyses a GTP molecule when the correct association has been found, thus acting as a Maxwell's demon. In this context we expect that energy-dependent degradation is the hallmark of «information gathering and utilising systems», typical of what life is [62]. Energy-dependent proteases are degradation machines composed of a sensor required for both substrate recognition and ATP-dependent selection for unfolding, and of a peptidase made of multiple subunits, required for substrate hydrolysis (see e.g. [63]). They use their ATPase subunits to choose between folded and unfolded substrates. Remarkably, acyldepsipeptides antibiotics uncouple degradation from energy consumption. They initiate proteolytic degradation without the control exerted by the ATPases. This unchecked activity, which demonstrates that the energy of ATP is not used in the very process of protein hydrolysis, leads to the inhibition of bacterial cell division and eventually cell death [64].

Such error-correcting routines are the norm in biological processes, and function as working analogues of Maxwell's demon, getting information and using it to reduce entropy at an exchange rate of kT ln2 joules per bit, rejecting errors to the environment at a high rate to maintain reliable operations. Thus, as can be seen in the buds of yeast, or creation of a mammalian egg, energy-dependent septins behave exactly as the demon behaves [62,65], preventing aged proteins to go in young cells, so that old cells can create a young offspring. This reflection is therefore at the core of what should be a renewed view of the process of ageing, especially during the scaling up of SB processes.

5. Final constraints on the reproducing chassis

5.1. The program, the chassis and the flywheel

Underlying what became SB is the hidden assumption that cells behave as electronic devices. The Biobrick™ Foundation, for example, creates, collects and standardises DNA "bricks" that are meant to be combined as the transistors that make the electronic circuits of microchips. Even if we do not ask the simple questions: in these chemical constructs where are the wires? and what is playing the role of electric current? this approach does as if the cell, where the circuit is implemented, behaved as a battery, providing the energy potential that drives the time-dependent behaviour of the electronic mimic. This view is fairly naive. Indeed, putting a genetic circuit in a cell will have a considerable impact on the way it manages its fluxes of matter and energy. Furthermore, while the engineering of genetic circuits is meant to provide highly non-linear behaviours (sensors, amplifiers, homeostatic regulation...) [55], the way cells adapt to the availability of carbon, nitrogen and sulfur sources is essentially linear [66,67].

This discrepancy matches the epistemological shift from physiology to molecular biology. Once pervasive, cell physiology is no longer fashionable. Jacques Monod's PhD thesis in bacterial physiology is one of the founding works of molecular biology. It derived phenomenological rules that displayed bacterial cell growth as a function of nutrients availability. Remarkably, it established that there is a hyperbolic relationship between growth rate and a growth-limiting carbon supply. This implied that there is a maximum growth rate, that was subsequently shown to be linearly dependent on the global gene expression capacity of the cell, mediated by the transcription/translation apparatus, essentially translation, with its considerable number of ribosomes [67]. This work predated the birth of molecular biology, while striking the end of physiological studies. With the advent of molecular biology, linearity was to be replaced by the intrinsic nonlinearity of the regulation of gene expression (also with a major contribution by Monod). Yet, growth requires that cells draw matter and energy either from some – limited – storage system or adjust their input and global biosynthesis, and experimental observations show that this is performed in a linear way, combining at least three types of syntheses for (i) a fixed set of housekeeping functions, (ii) the transcription/translation apparatus (dominated, by far, by ribosome biosynthesis) and (iii) other, environment-related functions [66,67].

Constructing cells that probe the presence of cues in the environment, based on highly non linear circuits – such as detecting explosives in mine fields [68] – should not pose difficult challenges, as there is considerable decoupling between the amount of energy required to emit light from a bioluminescent device and the cell energy and mass. In contrast, using the cell as a factory to generate chemicals, from antibiotics to biofuels, asks a

Q1 Please cite this article in press as: Danchin, A. Scaling up synthetic biology: Do not forget the chassis. FEBS Lett. (2012), doi:10.1016/j.febslet.2011.12.024

completely different question. Environmental bacteria that thrive in polluted environments have, in fact solved the riddle. Via horizontal gene transfer and much invention they carry over a considerable fraction of their genome as an extract of a cenome, the set of genes available in a particular environmental niche (biocenosis). And these genes behave in a way that is quite similar to what humans would like to construct. How is the coupling between the qualitative organisation of the gene network with the quantitative simplicity of physiology achieved? de Lorenzo and co-workers, using a Boolean formalism coupled to qualitative computation of general fluxes follow a path that may rewarding, identifying bottlenecks and high level organisation of regulatory circuits [55]. It is now time to go for experiments and find out what could be the flywheel that couples linear to non-linear behaviour: For example, interest in cyclic AMP and its catabolite regulation protein should be renewed. While it is known to be coupled to carbohydrate fluxes in *E. coli* (where it has been discovered), its targets are completely different in *Pseudomonas putida* [69]. Yet both the enzyme making cAMP and CRP are remarkably similar in both organisms. This shows that the features retained from their target is a not a particular property of a metabolic pathway, but a functional property, an information attached to the coupling between qualitative local organisation and quantitative global physiology.

5.2. Forgetting the chassis

If the cell is programmed to reproduce in such a way that it can recruit information from the environment, using its «Maxwell's demon's genes», then scaling up is doomed to fail, unless these genes can be harnessed to accumulate information toward the human goals, certainly a difficult task. This explains why, in general, evolving systems such as the one used to construct «*Escherichia chlori*» [70] cannot be stably maintained over time. There is however an alternative. The idea is to uncouple construction of biomass with using some of it to perform human-designed tasks. Panke and co-workers have paved the way in this domain [71,72]. The underlying idea is to create industrially useful pathway, first within cells. With proper adjustment of the corresponding expression patterns, they may reach a first level of functioning and make the cell produce the expected metabolite(s). However, despite all kinds of improvement in terms of promoter efficiency, codon usage bias etc, this will be at a level that is far too low to fit industrial requirements. The reason is that there is a conflict between the cell's agenda (which is to make a progeny in this particular environment) and the human agenda. In parallel with the required pathway many other begin to be triggered as soon as one attempts to improve the output. Here comes an interesting way to overcome this difficulty. The first task of the engineer, at this step, is to identify as many as possible of the «parasitic» pathways, that divert energy and biomass from the industrial goal. Once this is done (this will of course require a considerable amount of work), the idea is to tag all the genes involved with a short protease-specific target, that does not alter the overall functioning of the cell. Subsequently, a considerable volume of whole cell extract is produced and placed in a relevant reactor, together with a protease that will inactivate the parasitic enzymes. This way the yield of the required metabolite will be considerably increased [73]. Industrial processes can well be designed to follow this approach: 1/a cell factory; 2/a biochemical reactor.

6. Provisional conclusions

Synthetic Biology needs to take into account the chassis used to expressed the human-designed programs. It is therefore essential to extend the engineering reasoning to the chassis, and to uncover as many as possible of the unobtrusive functions that would be required to improved the yield of metabolic engineering. In this quest, it is essential to remember that scaling up implies that a given construct will be reproduced along many generations. Living organisms have an in-built ability to mobilise «Maxwell's demons genes» that will détourner the constructs from the human goals. Taming the demon should be the target of our present interests.

Acknowledgement

This work was supported by the FP7 European Union programme Microme KBBE-2007-3-2-08-222886-2 grant.

References

[1] Haseloff, J. and Ajioka, J. (2009) Synthetic biology: history, challenges and prospects. J. R. Soc. Interface 6 (Suppl. 4), S389–S391.
[2] Morange, M. (2009) A new revolution? The place of systems biology and synthetic biology in the history of biology. EMBO Rep. 10 (Suppl. 1), S50–S53.
[3] Yeh, B.J. and Lim, W.A. (2007) Synthetic biology: lessons from the history of synthetic organic chemistry. Nat. Chem. Biol. 3, 521–525.
[4] Tirard, S. (2009) Stephane Leduc et la biologie synthétique. Hist. Sci. Méd. 42, 67–72.
[5] Porcar, M., Danchin, A., de Lorenzo, V., Dos Santos, V.A., Krasnogor, N., Rasmussen, S. and Moya, A. (2011) The ten grand challenges of synthetic life. Syst. Synth. Biol. 5, 1–9.
[6] Lartigue, C., Glass, J.I., Alperovich, N., Pieper, R., Parmar, P.P., Hutchison 3rd, C.A., Smith, H.O. and Venter, J.C. (2007) Genome transplantation in bacteria: changing one species to another. Science 317, 632–638.
[7] Fantoni, G., Apreda, R. and Bonaccorsi, A. (2009) A theory of the constituent elements of functions in: Proceedings of the 17th International Conference on Engineering Design (ICED'09) (Norell Bergendahl, M., Grimheden, M., Leifer, L., Skogstad, P. and Lindemann, U., Eds.), pp. 179–190, Design Society, Stanford University, San Francisco, CA, USA.
[8] Dyson, F.J. (1985) Origins of life, Cambridge University Press, Cambridge, UK.
[9] Seely Jr., O., Feng, D.F., Smith, D.W., Sulzbach, D. and Doolittle, R.F. (1990) Construction of a facsimile data set for large genome sequence analysis. Genomics 8, 71–82.
[10] Danchin, A. (1989) Complete genome sequencing: future and prospects in: BAP 1988–1989 (Goffeau, A., Ed.), pp. 1–24, Commission of the European Communities, Brussels.
[11] Fraser, C.M. et al. (1995) The minimal gene complement of *Mycoplasma genitalium*. Science 270, 397–403.
[12] Lagesen, K., Ussery, D.W. and Wassenaar, T.M. (2010) Genome update: the 1000th genome – a cautionary tale. Microbiology 156, 603–608.
[13] Beller, H.R. et al. (2006) The genome sequence of the obligately chemolithoautotrophic, facultatively anaerobic bacterium *Thiobacillus denitrificans*. J. Bacteriol. 188, 1473–1488.
[14] Paul, S., Dutta, A., Bag, S.K., Das, S. and Dutta, C. (2010) Distinct, ecotype-specific genome and proteome signatures in the marine cyanobacteria *Prochlorococcus*. BMC Genomics 11, 103.
[15] Vincent, K.A., Cracknell, J.A., Clark, J.R., Ludwig, M., Lenz, O., Friedrich, B. and Armstrong, F.A. (2006) Electricity from low-level H2 in still air – an ultimate test for an oxygen tolerant hydrogenase. Chem. Commun. (Camb.) 5033, 5.
[16] Janssen, P.J. et al. (2010) The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. PLoS ONE 5, e10433.
[17] Fang, G., Rocha, E. and Danchin, A. (2005) How essential are nonessential genes? Mol. Biol. Evol. 22, 2147–2156.
[18] Fang, G., Rocha, E.P. and Danchin, A. (2008) Persistence drives gene clustering in bacterial genomes. BMC Genomics 9, 4.
[19] Danchin, A. (2009) Bacteria as computers making computers. FEMS Microbiol. Rev. 33, 3–26.
[20] Danchin, A. (1989) Homeotopic transformation and the origin of translation. Prog. Biophys. Mol. Biol. 54, 81–86.
[21] Ciampi, M.S. (2006) Rho-dependent terminators and transcription termination. Microbiology 152, 2515–2528.
[22] Olenin, S. and Ducrotoy, J.P. (2006) The concept of biotope in marine ecology and coastal management. Mar. Pollut. Bull. 53, 20–29.
[23] Tettelin, H. et al. (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial « pan-genome ». Proc. Natl. Acad. Sci. USA 102, 13950–13955.
[24] Touchon, M. et al. (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. PLoS Genet. 5, e1000344.
[25] Rohde, H. et al. (2011) N. Engl. J. Med..
[26] King, J., Griffin-Shea, R. and Fuller, M.T. (1980) Scaffolding proteins and the genetic control of virus shell assembly. Q. Rev. Biol. 55, 369–393.
[27] von Neumann, J. (1966) Theory of self-reproducing automata, University of Illinois Press, Urbana.

[28] Mingorance, J., Tamames, J. and Vicente, M. (2004) Genomic channeling in bacterial cell division. J. Mol. Recognit. 17, 481–487.
[29] Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W. and Duboule, D. (2011) The dynamic architecture of Hox gene clusters. Science 334, 222–225.
[30] Vicedo, M. (1992) The human genome project: towards an analysis of the empirical, ethical, and conceptual issues involved. Biol. Philos. 7, 255–278.
[31] Vendeville, A., Lariviere, D. and Fourmentin, E. (2011) An inventory of the bacterial macromolecular components and their spatial organization. FEMS Microbiol. Rev. 35, 395–414.
[32] Shea, T.B., Harrison, B. and Seaman, E. (1979) Electron microscopy of the morphogenesis of *Bacillus subtilis* bacteriophage SP3. J. Gen. Virol. 45, 515–518.
[33] Kunzler, P. and Berger, H. (1981) Structure of the scaffold in bacteriophage lambda preheads removal of the scaffold leads to a change of the prehead shell. J. Mol. Biol. 153, 961–978.
[34] Wood, W.B. (1980) Bacteriophage T4 morphogenesis as a model for assembly of subcellular structure. Q. Rev. Biol. 55, 353–367.
[35] Georgopoulos, C.P. and Hohn, B. (1978) Identification of a host protein necessary for bacteriophage morphogenesis (the groE gene product). Proc. Natl. Acad. Sci. USA 75, 131–135.
[36] Ellis, J. (1987) Proteins as molecular chaperones. Nature 328, 378–379.
[37] Hartl, F.U., Bracher, A. and Hayer-Hartl, M. (2011) Molecular chaperones in protein folding and proteostasis. Nature 475, 324–332.
[38] Shajani, Z., Sykes, M.T. and Williamson, J.R. (2011) Assembly of bacterial ribosomes. Annu. Rev. Biochem. 80, 501–526.
[39] Mogk, A., Huber, D. and Bukau, B. (2011) Integrating protein homeostasis strategies in prokaryotes. Cold Spring Harb. Perspect. Biol. 3.
[40] Gibson, D.G. et al. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome. Science 329, 52–56.
[41] Teif, V.B. and Bohinc, K. (2010) Condensed DNA: condensing the concepts. Prog. Biophys. Mol. Biol. 105, 208–222.
[42] Lentz, B.R. (2007) PEG as a tool to gain insight into membrane fusion. Eur. Biophys. J. 36, 315–326.
[43] Barth, M., Dederich, D. and Dedon, P. (2009) An improved method for large-scale preparation of negatively and positively supercoiled plasmid DNA. BioTechniques 47, 633–635.
[44] Nikaido, H. (2009) Multidrug resistance in bacteria. Annu. Rev. Biochem. 78, 119–146.
[45] Danchin, A. (2009) Cells need safety valves. BioEssays 31, 769–773.
[46] Marbach, A. and Bettenbrock, K. (2011) *lac* operon induction in *Escherichia coli*: systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. J. Biotechnol..
[47] Fang, M., Zeisberg, W.M., Condon, C., Ogryzko, V., Danchin, A. and Mechold, U. (2009) Degradation of nanoRNA is performed by multiple redundant RNases in *Bacillus subtilis*. Nucleic Acids Res. 37, 5114–5125.
[48] Plutarch (2001) Theseus in: Plutarch's Lives (Clough, A.H., Ed.), pp. 1–24, The Modern Library, New York.
[49] Schneider, T.D. (2010) A brief review of molecular information theory. Nano Commun. Netw. 1, 173–180.
[50] Greer, E.L. et al. (2011) Transgenerational epigenetic inheritance of longevity in *Caenorhabditis elegans*. Nature.
[51] Maxwell, J. (1871, reprinted 1902) Theory of heat, Longmans, Green and Co. London, New York, Bombay.
[52] Smoluchowski, M. (1914) Gültigkeitsgressen des zweiten Hauptsatzes der Wärmtheorie. In Vorträge über die kinetische Theorie der Materie und der Elektrizität gehalten in Göttingen auf Einladung der Kommission der Wolfskehlstiftung). Teubner, Leipzig und Berlin, pp. 87–122.
[53] Szilard, L. (1929) Uber die Entropieverminderung in einem thermodynamischen System be eingriffen intelligenter Wesen. Z. Phys. 53, 840–856.
[54] Landauer, R. (1961) Irreversibility and heat generation in the computing process. IBM J. Res. Dev. 3, 184–191.
[55] Silva-Rocha, R., Tamames, J., dos Santos, V.M. and de Lorenzo, V. (2011) The logicome of environmental bacteria: merging catabolic and regulatory events with Boolean formalisms. Environ. Microbiol. 13, 2389–2402.
[56] Bennett, C. (1988) Notes on the history of reversible computation. IBM J. Res. Dev. 44, 270–277.
[57] Younis, S. and Knight, T. (1993) Practical implementation of charge recovering asymptotically zero power CMOS in: 1993 symposium on Research on integrated systems (Borriello, G. and Ebeling, C., Eds.), pp. 234–250, MIT Press, Seattle, Washington.
[58] Shimizu, T., Matsuoka, Y. and Shirasawa, T. (2005) Biological significance of isoaspartate and its repair system. Biol. Pharm. Bull. 28, 1590–1596.
[59] Tarrago, L., Laugier, E. and Rey, P. (2009) Protein-repairing methionine sulfoxide reductases in photosynthetic organisms: gene organization, reduction mechanisms, and physiological roles. Mol. Plant 2, 202–217.
[60] Van Schaftingen, E., Collard, F., Wiame, E. and Veiga-da-Cunha, M. (2011) Enzymatic repair of Amadori products. Amino Acids.
[61] Hopfield, J.J. (1974) Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. Proc. Natl. Acad. Sci. USA 71, 4135–4139.
[62] Binder, P.M. and Danchin, A. (2011) Life's demons: information and order in biology. What subcellular machines gather and process the information necessary to sustain life? EMBO Rep..
[63] Chatterjee, I., Neumayer, D. and Herrmann, M. (2010) Senescence of staphylococci: using functional genomics to unravel the roles of ClpC ATPase during late stationary phase. Int. J. Med. Microbiol. 300, 130–136.
[64] Brotz-Oesterhelt, H. et al. (2005) Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. Nat. Med. 11, 1082–1087.
[65] Saarikangas, J. and Barral, Y. (2011) The emerging functions of septins in metazoans. EMBO Rep. 12, 1118–1126.
[66] Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z. and Hwa, T. (2010) Interdependence of cell growth and gene expression: origins and consequences. Science 330, 1099–1102.
[67] Scott, M. and Hwa, T. (2011) Bacterial growth laws and their applications. Curr. Opin. Biotechnol. 22, 559–565.
[68] de las Heras, A., Carreno, C.A. and de Lorenzo, V. (2008) Stable implantation of orthogonal sensor circuits in Gram-negative bacteria for environmental release. Environ. Microbiol. 10, 3305–3316.
[69] Milanesio, P., Arce-Rodriguez, A., Munoz, A., Calles, B. and de Lorenzo, V. (2011) Regulatory exaptation of the catabolite repression protein (Crp)-cAMP system in *Pseudomonas putida*. Environ. Microbiol. 13, 324–339.
[70] Marliere, P., Patrouix, J., Doring, V., Herdewijn, P., Tricot, S., Cruveiller, S., Bouzon, M. and Mutzel, R. (2011) Chemical evolution of a bacterium's genome. Angew. Chem., Int. Ed. Engl. 50, 7109–7114.
[71] Hold, C. and Panke, S. (2009) Towards the engineering of in vitro systems. J. R. Soc. Interface 6 (Suppl. 4), S507–S521.
[72] Bujara, M., Schumperli, M., Billerbeck, S., Heinemann, M. and Panke, S. (2010) Exploiting cell-free systems: Implementation and debugging of a system of biotransformations. Biotechnol. Bioeng. 106, 376–389.
[73] Hodgman, C.E. and Jewett, M.C. (2011) Cell-free synthetic biology: thinking outside the cell. Metab. Eng..