Organism Chemistry

DOI: 10.1002/anie.201103010

On the Road towards Chemically Modified Organisms Endowed with a Genetic Firewall**

*Carlos G. Acevedo-Rocha and Nediljko Budisa**

chemical evolution · nucleobases · synthetic biology · xenobiology

Living systems are chemical "machines" controlled by a genetic program with standardized compounds and controlled chemical processes. One of the greatest challenges for scientists is finding a way to expand the standard chemical repertoire of living cells. This can be achieved by directed artificial evolution of organisms with novel chemical compositions. The resultant cells are expected to be viable and robust for growth as well as replication for an unlimited time in genetic isolation from natural species. However, this is not a trivial task, as the organization of metabolic pathways and information processing as well as a standard set of macromolecules (nucleic acids, proteins, fatty acids) are common features of all currently known cells. These building blocks include the standard repertoire of 20 canonical amino acids for protein biosynthesis and, in the case when DNA is the genetic material, four types of canonical nucleotides: adenine (A), thymine (T), guanine (G), and cytosine (C). In addition, the universality of the genetic code enables the horizontal transfer of genes across biological taxa. As a consequence of this high degree of standardization and interconnectivity, fundamental chemical changes within living systems generally tend to be lethal.[1]

The breakthrough by Marlière et al. is the first successful attempt to cross the canonical/noncanonical chemical barrier by artificially evolving bacteria with a chlorinated DNA genome.[2] They showed that T could be "transliterated" to its noncanonical analogue 5-chorouracil (χ) in the genome of *Escherichia coli*, whose progeny retained the ability to use χ. This was achieved by extensively modifying the T biosynthetic pathway and by the successful selection of robust *E. coli* variants with the ability to grow on χ. The choice of χ is based on earlier reports that show that: 1) χ has long been known to be incorporated into DNA;[3a] 2) χ forms stable base pairs with A,[3b] and 3) χ is readily metabolized by the components of the T pathway in *E. coli*.[3c] Another advantage is that T is the only DNA base not used in RNA metabolism. In addition, Marlière et al. were already able to generate extremely oligotropic *E. coli* cells capable of growing on any spare T[4a] by using a method to evolve new life forms.[4b] The available intracellular thymine is normally metabolically converted into thymidine triphosphate (dTTP) as a building block for DNA. However, this pathway can also be used by χ (as a replacement for T) to generate chlorodeoxyuridine triphosphate (dχTP), which is enzymatically polymerized into chlorinated DNA.

Evolution of the *E. coli* strain THY1 (a derivative of the *E. coli* K12 strain MG1655) in the presence of χ under "harsh" and "gentle" experimental conditions yielded the CLU2 and CLU4 isolates, respectively. The experiment was stopped after the exclusive consumption of χ and concomitant generation of CLU2 and CLU4 (after 164 and 166 days, respectively); subsequent analysis of the DNA composition revealed 90% deoxychlorouridine (dχ) in both strains. The residual 10% deoxythimidine (dT) came from a rather unconventional source: In most tRNAs, the uracil (unmethylated T) at position 54 is methylated by the S-AdoMet-dependent 5-MeU-54-tRNA methyltransferase (encoded by the *trmA* gene). After knocking out this gene locus in the THY4 strain, the resulting THY5 variant was capable of growing on χ, thereby giving rise to the variant CLU5, whose genomic DNA contained only traces (1.6%) of dT. This is most probably due to other RNA- or DNA-modifying enzymes. Figure 1 presents the morphology of the native genomic DNA-containing strains as well as the one with a chlorinated genome.

When an organism is exposed to a novel environment, it can only adapt through massive modification of the enzymes and proteins that originally evolved to respond to quite different regimes.[5] Long-term cultivation experiments with fast-growing asexual bacterial cells—pioneered by Richard Lenski and co-workers—are the approach of choice to study these processes. In fact, they succeeded in cultivating up to 2000,[6a] 10000[6b], 20000,[6c] and 40000[6d] *E. coli* generations by serial dilution. However, this is technically tedious and time-consuming, because contaminations and discontinuous growth are difficult to avoid. In conventional continuous culture setups, on the other hand, undesired dilution-resistant adhesive strain variants usually stick to surfaces of the device.

[*] Dr. C. G. Acevedo-Rocha, Prof. Dr. N. Budisa
Biocatalysis Group, Institute of Chemistry
Berlin Institute of Technology/TU Berlin
Franklinstrasse 29, 10587 Berlin (Germany)
E-mail: budisa@biocat.tu-berlin.de
Homepage: http://www.biocat.tu-berlin.de
Dr. C. G. Acevedo-Rocha
Arbeitskreis Reetz, Institute of Chemistry
Philipps-Universität Marburg (Germany)

[**] We are indebted to Dr. Michael G. Hoesl and Dr. Lars Merkel for critical reading of this manuscript and numerous productive discussions on this subject.

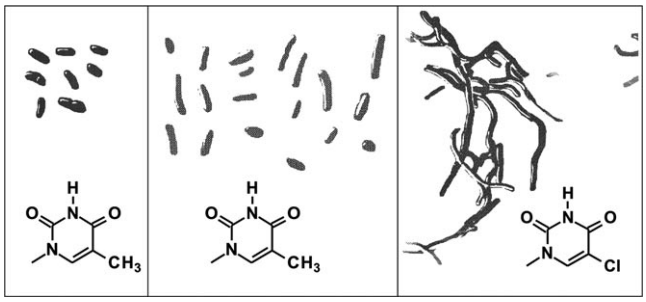

Figure 1. Marlière–Mutzel experiment—chemical evolution of a bacterial genome.[2] Graphical representation of the morphology of *E. coli* cells derived from the strains THY1 (left), THY5 (middle), and CLU5 (right). The graphics are based on micrographs kindly provided by R. Mutzel and P. Marlière. The original strain THY1 is rodlike, whereas the evolved 5-chlorouracil-resistant strain adopts two forms depending on the nutrient provided in the media: While the THY5 cells grown with thymine have a slightly elongated rodlike shape (middle), the growth of the same strain (CLU5) on 5-chlorouracil is characterized by a long filamentous shape of irregular density. The size increase is most probably due to changes in the dynamics of cell division.

Marlière and Mutzel elegantly circumvented these difficulties by inventing an ingenious device known as a "gene machine" or "genemat" (a kind of "automated Lenski" apparatus).[4b] The fluidic machine is equipped with twin chambers that undergo alternating sterilizations so that no adhesion variant escapes dilution and undesired selection. Two types of medium are alternated: a permissive/relaxing medium (with canonical metabolite) and a stress/growth-inhibiting one (with noncanonical metabolites). The working threshold for the culture is set to 10^9 cells per milliliter. The growth cycles are automatically controlled by a simple algorithm: if the optical density of the culture is below a certain threshold value, relaxing media is delivered, and vice versa.

Marlière et al.,[2] by using the modest "genetic engine" setup, were able to gradually adapt *E. coli* from a relaxing medium containing T to a "stressing" medium harboring the χ. Both the CLU2 and CLU4 strains, however, retained their ability to proliferate indefinitely on T at a similar growth rate as on χ. By subsequently growing the χ-adapted CLU2 and CLU4 strains in the presence of T, they were able to generate the strains THY2 and THY4, respectively. Nevertheless, the most interesting finding was that the strains THY2 and THY4 were able to grow again on χ-containing media without delaying their adaptation, which is in contrast to the CLU2 strain when growing solely in the presence of T.[2] Most importantly, the ability to build chlorinated DNA was encoded in the genome of the strains THY2 and THY4, provided their genome was highly mutated: DNA sequence analysis revealed 1514 substitution mutations and a deletion of at least 150 kilobases for THY2, as well as several mutations and a 20% rearrangement of the genome (846 kilobases) for THY4. This experiment demonstrates that the accumulation of different types of mutations required the editing of a remarkably large portion of the genome. An application of these strains could be to discover new biosynthetic properties of the mutated enzymes in the related pathways or to further tailor the evolved stains as whole-cell catalysts for specific needs.

The in vivo delivery of novel or nonstandard processes and compounds is rather limited as it usually allows only small/negligible variations in the basic design of the living cells.[1] Consequently, it may be thought that reprogramming cells with synthetic ingredients would not compete with the standardized chemistry and functional diversity found in the natural evolution of organisms. However, the combination of experimental evolution,[6,7] chemical evolution,[8] and natural genetic engineering (genetic mutations, recombination, horizontal gene transfer, etc.)[9] is a powerful approach to generate truly artificial cells with new functions not found in nature. It has been foreseen that unanticipated life forms and properties could emerge from cell populations with fully integrated novel chemical functionalities.[1a,10a] However, it should be kept in mind that the evolved bacteria containing a chlorinated genome still have the ability to grow on T. Hence, the next step will be to achieve a complete metabolic/genetic isolation, that is, full dependence on χ with concomitant metabolic indifference towards T. In the same vein, the use of noncanonical amino acids and nucleobases should be only the first step in changing chemical compositions of cells.[1] Thereafter, the functional features of living systems could be further diversified by acquisition of novel chemical functionalities through artificial evolution of pre-existing cells from the "old" natural world. An important task for chemical synthetic biology would be the artificial evolution of chemically diverse viable and robust cells that grow and replicate for an unlimited time in isolation from natural species. Complete genetic isolation, which is only possible with cells containing "xeno-DNA" as genetic material or alternative/different genetic codes, should prevent horizontal gene transfer between species. In this way, a genetic firewall against the natural DNA-based world could be established. The biosafety and biosecurity aspects of these possibilities have been extensively elaborated just recently.[10]

It is anticipated that experiments aimed at changing the interpretation of the genetic code (e.g. through codon-emancipation for novel amino acid assignment) will be performed in the future by using "genetic engines". Directed evolution of bacterial strains should likewise include non-canonical DNA base pairs[11] to increase the coding capacity of cells. The ultimate goal is an integrative approach that will include directed evolution of artificial cells with expanded coding capacities through novel genetic codes and de novo metabolic circuits, thereby providing living systems not only with new chemical compositions but also with emergent properties not originating in natural evolution.[1] These developments will thus open the door to a parallel biological world.

Received: May 2, 2011
Published online: June 24, 2011

[1] a) N. Budisa, *Angew. Chem.* **2004**, *116*, 6586–6624; *Angew. Chem. Int. Ed.* **2004**, *43*, 6426–6463; b) N. Budisa, *Engineering the genetic code*, Wiley-VCH, Weinheim, **2005**; c) M. Hoesl, N.

Budisa, *Angew. Chem.* **2011**, *123*, 2948 – 2955; *Angew. Chem. Int. Ed.* **2011**, *50*, 2896 – 2902.
[2] P. Marlière, J. Patrouix, V. Döring, P. Herdewijn, S. Tricot, S. Cruveiller, M. Bouzon, R. Mutzel, *Angew. Chem.* **2011**, DOI: 10.1002/ange.201100535; *Angew. Chem. Int. Ed.* **2011**, DOI: 10.1002/anie.201100535.
[3] a) D. B. Dunn, J. D. Smith, *Biochem. J.* **1957**, *67*, 494 – 506; b) J. A. Theruvathu, C. H. Kim, D. K. Rogstad, J. W. Neidigh, L. C. Sowers, *Biochemistry* **2009**, *48*, 7539; c) A. Munch-Petersen: *Metabolism of nucleotides, nucleosides and nucleobases in microorganisms*, Academic Press, London, **1983**.
[4] a) V. de Crécy-Lagard, J. Bellalou, M. Bouzon, R. Mutzel, P. Marlière, *BMC Biotechnol.* **2001**, *1*, e10; b) P. Marlière, R. Mutzel, Patent DE 198 56 136 C2, **2002**.
[5] a) G. Bell, *Selection: the Mechanism of Evolution*, Chapman & Hall, New York, **1996**; b) J. W. Drake, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7160 – 7164.
[6] a) R. E. Lenski, M. R. Rose, S. C. Simpson, S. C. Tadler, *Am. Nat.* **1991**, *138*, 1315 – 1341; b) R. E. Lenski, M. Travisano, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6806 – 6814; c) R. E. Lenski, C. L. Winkworth, M. A. Riley, *J. Mol. Evol.* **2003**, *56*, 498 – 508; d) J. E. Barrick, D. S. Yu, S. H. Yoon, H. Jeong, T. K. Oh, D. Schneider, R. E. Lenski, J. F. Kim, *Nature* **2009**, *461*, 1243 – 1247.
[7] a) S. F. Elena, R. E. Lenski, *Nat. Rev. Genet.* **2003**, *4*, 457 – 469; b) N. Philippe, E. Crozat, R. E. Lenski, D. Schneider, *Bioessays* **2007**, *29*, 846 – 860; c) A. Buckling, R. Craig Maclean, M. A. Brockhurst, N. Colegrave, *Nature* **2009**, *457*, 824 – 829.
[8] S. J. Wrenn, P. B. Harbury, *Annu. Rev. Biochem.* **2007**, *76*, 331 – 349.
[9] J. A. Shapiro, *Mob. DNA* **2010**, *1*, 4.
[10] a) P. Marlière, *Syst. Synth. Biol.* **2009**, *3*, 77 – 84; b) P. Herdewijn, P. Marlière, *Chem. Biodiversity* **2009**, *6*, 791 – 808; c) M. Schmidt, *BioEssays* **2010**, *32*, 322 – 331.
[11] a) S. A. Benner, *Science* **2004**, *306*, 625 – 626; b) A. T. Krueger, E. T. Kool, *Chem. Biol.* **2009**, *16*, 242 – 248.

