

## Chemical Evolution of a Bacterium's Genome\*\*

Philippe Marlière, Julien Patrouix, Volker Döring, Piet Herdewijn, Sabine Tricot, Stéphane Cruveiller, Madeleine Bouzon, and Rupert Mutzel\*

We set out to develop a generic technology for evolving the chemical constitution of microbial populations by using the simplest possible algorithm. Extant living cells polymerize a restricted set of nucleic acid precursors, namely, four nucleoside triphosphates (UTP, CTP, ATP, GTP) and four deoxynucleoside triphosphates (dTTP, dCTP, dATP, dGTP).<sup>[1]</sup> Synthetic analogues, such as 5-halogenopyrimidines, 7-deazapurines, and 8-azapurines, are known to partially replace canonical bases in cellular RNA and DNA, yet were never demonstrated to sustain unlimited self-reproduction of an organism through complete genome or transcriptome substitution.<sup>[2]</sup> A hamster cell line serially adapted to grow in the presence of bromodeoxyuridine, while dTMP synthesis was inhibited with aminopterin, has been reported to harbor DNA highly enriched in bromouracil over thymine.<sup>[3,4]</sup> However, the significance of these findings could not be ascertained owing to the absence of a direct physical measurement of the base composition of the DNA and the absence of an assay of thymidylate biosynthesis, as well as the likely presence of metabolic components, such as nucleotides in the complex growth medium of the cells. Only certain DNA viruses are known to have undergone full transliteration of a canonical base through the biosynthesis of a noncanonical nucleoside triphosphate, for example, hydroxymethylcytosine in the T4 bacteriophage, presumably to counteract the restriction

enzymes of their bacterial hosts.<sup>[5]</sup> When Weiss and co-workers attempted to substitute thymine in the DNA of *Escherichia coli* with uracil, over 90% replacement was reached, but further growth was prevented.<sup>[6]</sup> Genome-scale transliteration has apparently not evolved in any known living cell, possibly owing to a chemical barrier that natural biodiversity cannot overcome. Our experimental plan consisted of the combination of tight metabolic selection with the long-term automated cultivation of fast-growing asexual bacterial populations to change a canonical DNA base for a chemical ersatz.

The cultivation setup was elaborated from the GM3 fluidic format (Figure 1), which features the cyclic transfer of the culture between twin growth chambers that alternately undergo sterilization.<sup>[7]</sup> This cycle ensures that no internal surface of the device is spared from transient periodic cleansing with a sterilizing agent (5M sodium hydroxide), and therefore that no cultivated variant can escape dilution and selection for faster growth through the formation of biofilms.<sup>[8a]</sup> The active elimination of biofilms (wall growth) has proved critical for reprogramming and improving the metabolism of microbial populations.<sup>[8b]</sup>

The GM3 cultivation device was connected to two nutrient reservoirs of different composition: a relaxing medium R that contains the canonical nutrient and a stressing medium S that contains the ersatz nutrient. Liquid pulses of defined volume are sent at regular intervals of time from these reservoirs to the culture, which is kept at a constant volume. Depending upon the state of the adapting cells, as measured by turbidity recording of the population density, the culture periodically receives a pulse of fixed volume of either medium R (if the population density falls below a fixed threshold) or medium S (if the density is higher than or equal to the threshold). Successive pulses thus renew the culture at a fixed dilution rate with a nutrient-medium flow whose composition varies with respect to the growth response of the population in such a way that the lowest tolerable concentration of canonical nutrient is automatically maintained over passing generations.

We designate this mode of operation as the conditional pulse-feed regime. It qualifies as a simplified and generalized version of a method pioneered by Oliver.<sup>[9,10]</sup> Mutations that confer a lower requirement for the canonical nutrient or a higher survival rate under starvation are expected to accumulate in the genome of the adapting population.<sup>[11]</sup> No attempt was made to implement a finer regulation of differential nutrient supply than the coarse-grained control by medium-switch pulse feed described above. We thus relied on the robustness of biochemical machineries and their evolution to dampen oscillations.

[\*] Dr. P. Marlière  
Heurisko USA Inc., Delaware (USA)

J. Patrouix, Dr. V. Döring, S. Tricot, Dr. S. Cruveiller, Dr. M. Bouzon  
CEA, DSV, IG, Genoscope, Evry (France)

Prof. Dr. P. Herdewijn  
Katholieke Universiteit Leuven, Rega Institute for Medical Research  
Laboratory of Medicinal Chemistry, Leuven (Belgium)

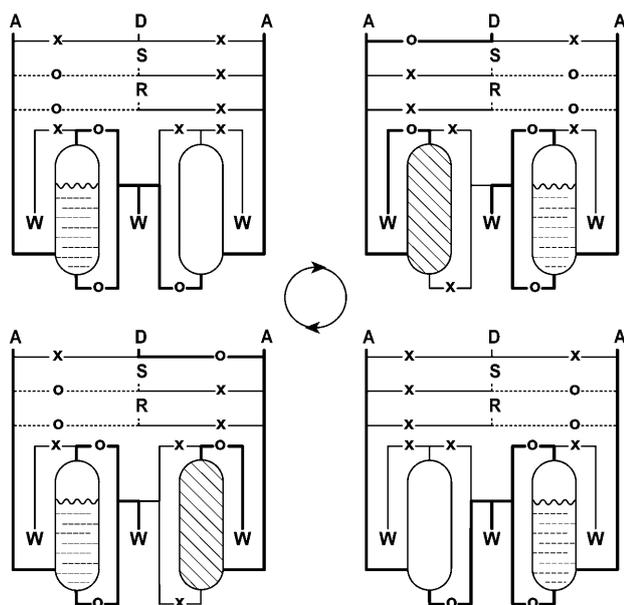
S. Tricot, Dr. S. Cruveiller  
CNRS, UMR Génomique Métabolique  
and  
Université Evry Val d'Essonne (France)

Prof. Dr. R. Mutzel  
Fachbereich Biologie, Chemie, Pharmazie  
Freie Universität Berlin  
Königin-Luise-Strasse 12–16, 14195 Berlin (Germany)  
Fax: (+30) 8385-7773  
E-mail: rmutzel@zedat.fu-berlin.de

[\*\*] We are grateful to Jean Weissenbach for his constant support and encouragement. We also thank Sophie Tuffet for DNA-composition analysis; Julie Poulain, Karine Labadie, Valérie Barbe, and Béatrice Chane-Woon-Ming for genome sequencing and annotation; Valérie Delmas for strain construction; Isabelle Boko and Angela Lahrz for technical assistance; and Susan Cure, Sven Panke, and Phil Holliger for improving the manuscript.

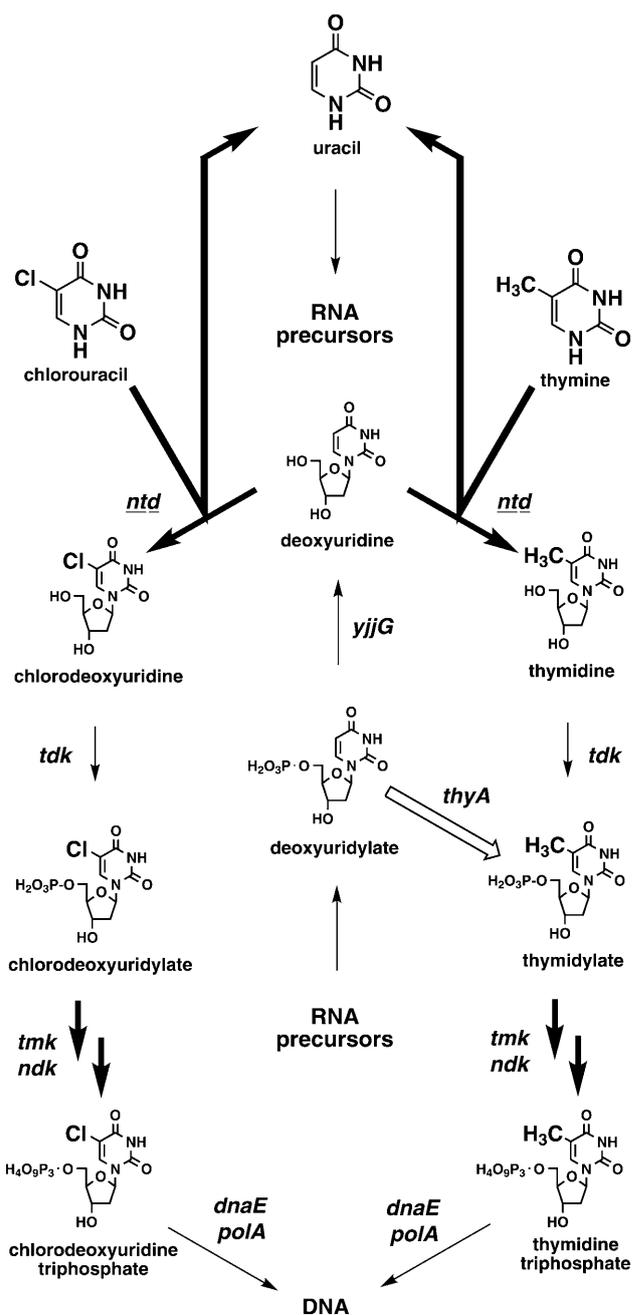


Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201100535>.



**Figure 1.** Cultivation fluidics with twin growth chambers and alternative nutrient media. Only four configurations of the symmetrical device are shown of the 64 successive steps carried out in every cycle of 12 h. Fluids flow from pressurized reservoirs containing air (A), rinsing water (not shown), 5 M sodium hydroxide (D), a relaxing nutrient medium containing thymine (R), and a stressing nutrient medium containing chlorouracil but no thymine (S) through two cultivation chambers down to waste at atmospheric pressure (W). Valves are indicated by x and o for the closed and open states, respectively. Constant and intermittent flows in tubing are shown by bold and dotted lines, respectively. Dashes in chambers indicate the bacterial population and oblique stripes, sodium hydroxide. The opening of medium valves is conditional on the density of the bacterial culture as measured by diode–photocell forks mounted on each chamber (not shown). At regular time intervals, a pulse of S medium is delivered if the density exceeds a certain threshold; a pulse of R medium is delivered otherwise. The cultivation process is carried out as shown in succession from the top left (clockwise): cultivation in the left chamber and emptying of the rinsed right chamber; culture transfer then cultivation in the right chamber and purging of the left chamber; cultivation in the right chamber and emptying of rinsed left chamber; culture transfer then cultivation in the left chamber and purging of the right chamber.

Thymine is the only nucleobase specifically present in DNA, and its metabolism is cleanly disentangled from RNA biosynthesis; therefore, the incorporation of thymine analogues *in vivo* can be manipulated more easily than the replacement of other nucleobases.<sup>[1]</sup> In *E. coli*, the biosynthesis of thymine nucleotides can be disabled by simply disrupting the *thyA* gene for thymidylate synthase, which produces deoxythymidine monophosphate (thymidylate, dTMP) from deoxyuridine monophosphate (deoxyuridylate, dUMP) and methylenetetrahydrofolate, with the release of dihydrofolate (Figure 2; see Figure S1a in the Supporting Information for details).<sup>[2]</sup> Thymine starvation in *E. coli* and other bacteria results in the rapid loss of viable cell titer, a well-studied response known as thymineless death.<sup>[12,13]</sup> Exogenous thymine and thymidine can be utilized for the synthesis of deoxythymidine triphosphate (dTTP) and to rescue *thyA*



**Figure 2.** Parallel metabolic conversion of thymine (5-methyluracil) and 5-chlorouracil into DNA nucleotides in reprogrammed *E. coli* strains. The two pyrimidine bases are condensed with the deoxyribose moiety of deoxyuridine by the same enzyme, nucleoside deoxyribosyltransferase encoded by the *ntd* gene from *Lactobacillus leichmannii*. The two resulting deoxynucleosides chlorodeoxyuridine (d<sub>x</sub>) and thymidine (dT) are then channeled to the corresponding triphosphates (d<sub>x</sub>TP and dTTP, respectively) by the enzymes thymidine kinase, thymidylate kinase, and nucleoside diphosphate kinase, which are encoded by the genes *tdk*, *tmk*, and *ndk*. The two main DNA polymerases, encoded by *dnaE* and *polA*, then catalyze the templated incorporation of the competing triphosphates into DNA. Reversible steps in the network are indicated by bold arrows. A ghost arrow shows the step disabled by deletion of the *thyA* gene for thymidylate synthase, the enzyme that forms the thymine moiety in the deoxynucleotide pool of wild-type *E. coli* by converting deoxyuridylate (dUMP) into thymidylate (dTMP) and concomitantly methylenetetrahydrofolate into dihydrofolate.

mutants by a salvage pathway which converts thymine into dTMP through the action of pyrimidine nucleoside phosphorylases (*udp* and *deoA* gene products) and thymidine kinase (*tdk* gene product; see Figure S1a).

It has long been known that the whole series of 5-halogenopyrimidines 5-fluoro-, 5-chloro-, 5-bromo-, and 5-iodouracil can be incorporated into nucleic acids.<sup>[14]</sup> We chose chlorouracil ( $\chi$ ) as the most promising candidate for DNA-base translocation (see Figure S1b) because: 1) it closely resembles thymine in structural studies of synthetic DNA duplexes, which suggests that the stability of the A: $\chi$  pair is close to that of A:T;<sup>[15]</sup> 2) it is readily converted into the chlorodeoxyuridine nucleoside and the d $\chi$ MP, d $\chi$ DP, and d $\chi$ TP nucleotides by nucleoside phosphorylases, thymidine kinase, thymidylate kinase, and nucleoside diphosphate kinase, respectively;<sup>[2]</sup> 3) unlike fluorouracil, it is less liable to elevated tautomerism, which causes ambiguous pairing with guanine as well as adenine;<sup>[15]</sup> 4) unlike the bromo and iodo analogues, it is not reduced to uracil at the redox potential of the cytoplasm.<sup>[16]</sup>

Chlorine is present in numerous natural compounds, especially secondary metabolites of marine organisms, and a number of enzymatic mechanisms that enable the formation of C–Cl bonds have been elucidated.<sup>[17]</sup> Apparently, it has not been reported to occur in any natural nucleic acid building block so far. Literature on the response of *E. coli* to chlorouracil incorporation is scarcer than that on bromouracil substitution. The effects for both analogues have been described as very similar and encompass filamentation, cessation of growth, low mortality, and mutagenesis.<sup>[16,18]</sup>

The strain THY1 that was subjected to 5-chlorouracil translocation is a derivative of wild-type *E. coli* K12 strain MG1655.<sup>[19]</sup> Its construction involved consecutive deletions of the genes *thyA* and *udp* and the *deoCABD* operon from the MG1655 chromosome, and the insertion of a P15A plasmid carrying the *Lactobacillus leichmannii* gene *ntd*, which encodes nucleoside deoxyribosyltransferase (see Supporting Information). This enzyme catalyzes the reversible conversion of thymine (T) into thymidine (dT) or 5-chlorouracil ( $\chi$ ) into 5-chlorodeoxyuridine (d $\chi$ ) by the use of deoxyuridine (dU) as a cosubstrate:<sup>[20,21]</sup>  $\chi + dU \rightleftharpoons d\chi + U$  and  $T + dU \rightleftharpoons dT + U$ , and enables growth in the absence of thymidine phosphorylase (*deoA*) and uridine phosphorylase (*udp*). Thymidine or 5-chlorodeoxyuridine can then be irreversibly channeled to DNA biosynthesis to provide d $\chi$ TP and dTTP as competing substrates for incorporation by DNA polymerases in response to adenine in DNA templates (see Figure 2). The dU cosubstrate of these reactions (2'-deoxyuridine) originates from dUTP by the successive and irreversible action of the *dut* and *yjg* gene products,<sup>[2,22]</sup> and the uracil coproduct (U) is recycled to RNA biosynthesis by the successive and irreversible action of the *upp*, *pyrH*, and *ndk* gene products. This integrated metabolic network conferred to the strain THY1 the ability to grow in the presence of thymine at a concentration as low as 1  $\mu$ M, a trait that did not revert even when vast populations were starved of thymine for long durations.<sup>[8]</sup> Chlorouracil could not sustain the proliferation of THY1 in liquid or on solid nutrient medium.

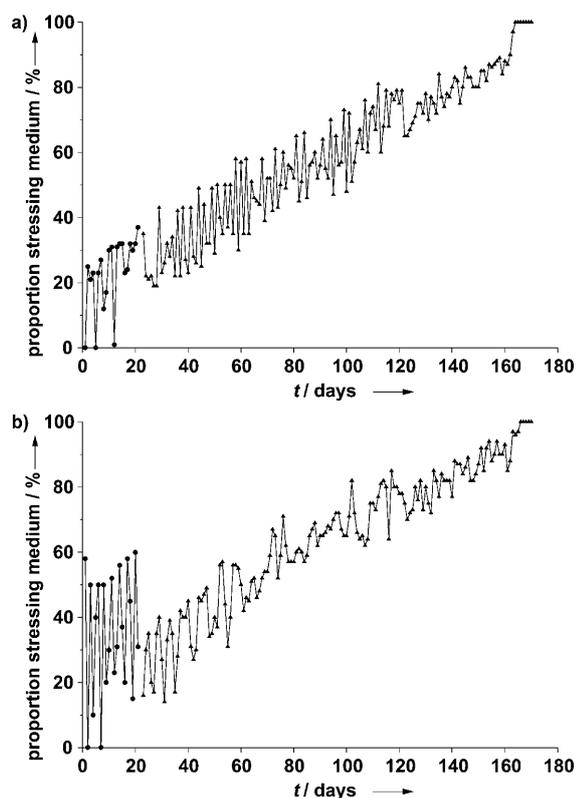
A population of THY1 cells was inoculated into a GM3 device connected to two nutrient reservoirs, the first containing thymine (10  $\mu$ M; relaxing medium) and the second chlorouracil (10  $\mu$ M; stressing medium). A generation time of 2 h was imposed under the conditional pulse-feed regime in mineral glucose medium at 37°C through the injection of nutrient pulses of either relaxing or stressing medium every 10 min. The composition of nutrient pulses was determined by the culture turbidity relative to a threshold of OD<sub>880</sub> = 1 (about 10<sup>9</sup> cells per milliliter). In parallel, a culture was initiated with an imposed generation time of 4 h by halving the volume of individual pulses delivered at the same frequency. Figure 3 shows the evolutionary kinetics for the two experiments.

Both cultures showed massive and sustained oscillations of cell density, yet displayed a trend of increasing consumption of the stressing medium (containing chlorouracil and lacking thymine). After 23 days, the resilience toward chlorouracil of adapting bacteria in both cultures, as established in batch cultures inoculated from samples, was judged sufficient for the application of a harsher regime to the two populations evolving from THY1. Both cultures were therefore reconnected to nutrient reservoirs containing a lower pyrimidine concentration: 3  $\mu$ M thymine with 3  $\mu$ M chlorouracil as the relaxing medium, and 3  $\mu$ M chlorouracil as the stressing medium.

In this way, adapting cells had to adjust to a constant concentration of the analogue and did find relief in lower amounts of the canonical base. Adaptation smoothly ensued under this harsher regime until only the stressing medium was consumed after 141 further days of cultivation at a generation time of 2 h (Figure 3a) and 143 further days at a generation time of 4 h (Figure 3b); these time periods correspond to a total of about 2000 and 1000 generations since divergence from their common progenitor THY1.

One clone was reisolated from each culture on solid nutrient medium for thorough study: CLU2 from the GM3 device set at a generation time of 2 h and CLU4 from the device set at a generation time of 4 h. Both strains showed an absolute growth requirement for chlorouracil. Reacclimatization of CLU2 and CLU4 on thymine in batch cultures through serial transfer was possible, albeit after an adaptive lag phase in the case of CLU2 (see Figure S2). The resulting strains, THY2 and THY4, could proliferate indefinitely with thymine (see Table S1 in the Supporting Information for strain construction and nomenclature). Similar growth rates were observed whether thymine or chlorouracil was used, for CLU2 and for CLU4 (data not shown). In turn, the thymine-containing strains THY2 and THY4 could be grown again with chlorouracil without any delay of adaptation. This result proves that the ability to construct DNA with chlorouracil was inheritable and encoded in the genome of the adapted strains. By comparison, uracil could not substitute for chlorouracil or thymine in either strain.

The viability of CLU2 and CLU4 was much reduced during the stationary phase, a likely consequence of the fact that selection for chlorouracil usage was enforced by a cultivation regime of permanent proliferation. As compared with their common THY1 progenitor, the bacterial cells from



**Figure 3.** Evolutionary kinetics of thymine replacement with chlorouracil in *E. coli*. The daily fraction of thymineless medium pulses (proportion stressing medium) infused into the culture is plotted as a function of time (elapsed cultivation days). The thymine-auxotrophic strain THY1 was adapted to grow under the medium-exchange pulse-feed regime. At regular time intervals of 10 min, the optical density of the culture at 880 nm was measured and compared to a fixed threshold ( $OD = 1$ ) corresponding to a bacterial density of about  $10^9$  cells per milliliter. When the measured optical density ( $OD$ ) exceeded the threshold, a pulse of stressing medium, lacking thymine, was infused into the culture; otherwise, a pulse of relaxing medium, containing thymine, was infused. The culture volume was kept constant at 16 mL, and the generation time was set by the volume of the pulses infused every 10 min. a) Continuous culture with a generation time of 2 h as set by infused pulses with a volume of 950  $\mu\text{L}$  (the CLU2 isolate was obtained after 164 days). b) Continuous culture with a generation time of 4 h as set by infused pulses with a volume of 460  $\mu\text{L}$  (the CLU4 isolate was obtained after 166 days). In both plots, circles indicate a first adaptation stage with chlorouracil (10  $\mu\text{M}$ ) in the stressing medium and thymine (10  $\mu\text{M}$ ) in the relaxing medium; triangles indicate a second harsher stage with chlorouracil (3  $\mu\text{M}$ ) in the stressing medium and chlorouracil (3  $\mu\text{M}$ ) together with thymine (3  $\mu\text{M}$ ) in the relaxing medium.

the two cultures grown in parallel displayed a rather uniform distribution of elongated rods different from the very long filaments of irregular density which could be observed under the microscope earlier in the evolutionary process.

Determination of the DNA composition of the strains grown with either 5-chlorouracil (CLU2 and CLU4) or thymine (THY1) through DNA extraction followed by enzymatic hydrolysis and the HPLC fractionation of deoxynucleosides revealed the massive incorporation of chloro-deoxyuridine in CLU2 and CLU4. The detected dT fraction was reduced to about 10% of the total d $\chi$  + dT content in

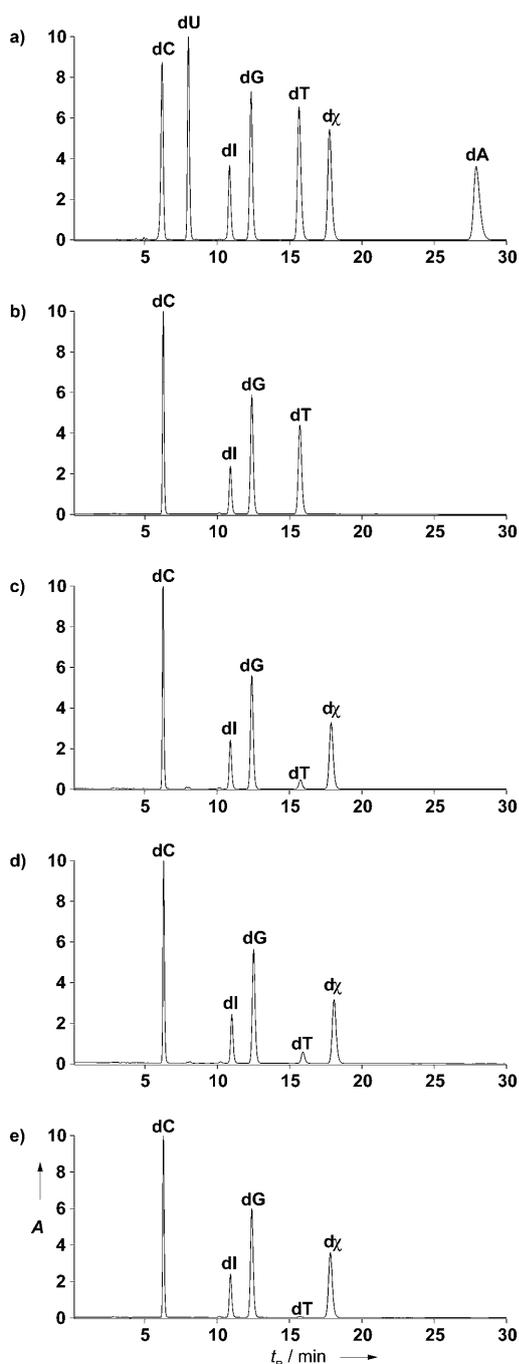
both strains (Figure 4). Traces of deoxyuridine were also detected, but they did not exceed the marginal amounts found in the DNA of the progenitor strain THY1.

Part of the remaining thymidine might originate through undefined recycling pathways from position 54 of tRNAs. This position is universally occupied by uridine and converted posttranscriptionally into ribothymidine in bacteria and eukaryotes.<sup>[23]</sup> The possible contribution of such alternative metabolic sources of thymine was addressed by disrupting the *trmA* gene, which in *E. coli* encodes *S*-adenosylmethionine-dependent U54 tRNA methyltransferase.<sup>[24]</sup> Strain THY2 proved refractory to allelic replacement either by P1 transduction or by the Wanner procedure.<sup>[25]</sup> By contrast, the *trmA* gene could be disrupted in strain THY4 through P1 transduction, which yielded the isolate THY5. The subsequent growth of this thymine-containing strain on chlorouracil yielded strain CLU5 (see Table S1 for strain construction and nomenclature). No phenotypic trait (cell shape, growth with various nutrients, temperature and antibiotic sensitivity) could be found that distinguished CLU5 from its progenitor CLU4. However, the deoxynucleoside-fractionation profile of CLU5 showed a drop in thymidine concentration to barely detectable traces (about 1.5% of the total d $\chi$  + dT content). This result demonstrated the contribution of the U54 tRNA methyltransferase activity to thymine production (Figure 4).

Mass spectrometric analysis confirmed the presence of chloro-2'-deoxyuridine ( $m/z$  263.0416) in the DNA of the strains CLU2, CLU4, and CLU5; this residue was absent in THY1. The residual thymine content, as quantified by using calibrating curves for thymidine and 5-chloro-2'-deoxyuridine, amounted to 8.8, 10.8, and 1.6% for CLU2, CLU4, and CLU5, respectively (see Table S2). Other modification enzymes present in *E. coli* that methylate the C5 position of uracil<sup>[26]</sup> and cytosine<sup>[27]</sup> in rRNA and tRNA molecules or of cytosine in DNA<sup>[28]</sup> presumably account for the remaining thymine in the genome of CLU5. All these *E. coli* enzymes are known to use the cosubstrate *S*-adenosylmethionine as a methyl donor, as does the *trmA* gene product.<sup>[26,27]</sup>

To determine mutations fixed in the genomes during adaptation to chlorouracil usage, we sequenced the DNA from the strains THY1, THY2, and THY4. The two evolved strains underwent quite dissimilar distortions of their genomes: THY2 had accumulated numerous base substitutions and THY4 chromosome rearrangements (Table 1; see Tables S3–S6 for details).

From a total of 1514 base substitutions accumulated in the genome of THY2, 1023 were A:T to G:C transitions, as compared with 479 G:C to A:T transitions. This result suggests that chlorouracil is prone to mispairing with guanine, and that such mispairing was more frequent when chlorouracil was in the template than when it was in the triphosphate substrate.<sup>[29]</sup> The mispairing could be caused by changes in the hydrogen-bond pattern and influenced by parameters such as  $pK_a$  value, dipole moment, and hydration.<sup>[30]</sup> Of these possibilities, formation of the anionic enol tautomer can be considered as a major factor that influences mispairing. The imino proton in chlorouracil ( $pK_a = 7.9$ ) is known to be more acidic than that in thymine ( $pK_a = 9.7$ ) but less acidic than in fluorouracil ( $pK_a = 7.7$ ),<sup>[15]</sup> which suggests that tautomerism



**Figure 4.** Deoxynucleoside composition of genomic DNA from chlorouracil-adapted bacterial cells. DNA extracted from cultures of THY1 and its chlorouracil-dependent derivatives CLU2, CLU4, and CLU5 was digested with nuclease P1, dephosphorylated, and deaminated. The deoxynucleosides were injected into an Uptisphere 5 ODB HPLC column and separated in aqueous buffer (12.5 mM citric acid, 25 mM Na acetate, 30 mM NaOH, pH 5.3) with 10% methanol at a flow rate of 0.8 mL min<sup>-1</sup>. A mixture of standard deoxynucleosides was separated following the same procedure. The elution profiles ( $t_R$  = retention time) of the deoxynucleosides were recorded on the basis of their absorbance at 272 nm: a) standards; b) THY1; c) CLU2; d) CLU4; e) CLU5. Deoxyadenosine (dA) was completely converted into dI by deamination during sample preparation (see the Supporting Information). dC, deoxycytidine; dU, deoxyuridine; dI, deoxyinosine; dG, deoxyguanosine; dT, thymidine; dχ, 5-chlorodeoxyuridine.

**Table 1:** Mutations in *E. coli* strains THY2 and THY4.<sup>[a]</sup>

| Mutation            | THY2 | THY4   |
|---------------------|------|--|
| substitutions       |      |  |
| A:T → G:C           | 1023 | 45   |
| G:C → A:T           | 479  | 53   |
| A:T → C:G           | 2    | 7  |
| A:T → T:A           | 5    | 1  |
| C:G → A:T           | 4    | 7  |
| C:G → G:C           | 1    | 13   |
| total substitutions | 1514 | 126  |
| Indels              |      |  |
|                     | 23   | 2  |
| rearrangements      |      |  |
| Inversion           |      | 84 6186<br>( <i>hns-asmA</i> )   |
| Deletions           |      |  |
|                     |      | 6650<br>(6 genes: <i>insF-2-yaiW</i> )<br>27 015<br>(18 genes: <i>rutE-serC</i> )<br>15 0814<br>(113 genes: <i>uspF-ydeP</i> ) |
|                     |      | 102164<br>(83 genes: <i>hrpA-ydeP</i> )<br>38 117<br>(34 genes: <i>elaB-yfcD</i> )   |

[a] Genomes of strains THY1, THY2, and THY4 were sequenced by using both the 454-Titanium and Solexa sequencing technologies. Mutations found in the genomes of THY2 and THY4 as compared to THY1 are summarized. The deletion length, the corresponding chromosomal regions, and the number of genes affected are indicated. A full list of point mutations and indels can be found in Tables S3–S6 of the Supporting Information.

of chlorouracil is intermediate between thymine and fluoro-uracil. Even though chlorouracil is not notorious as a mutagen and was shown to transiently and nonlethally substitute most of the thymine in the DNA of eukaryotic and bacterial cells,<sup>[14]</sup> a slight bias might nevertheless lead to a strong drift in G + C content over passing generations. By contrast, base substitutions in THY4, totaling 126 changes, were much less abundant than in THY2, and no preference for A:T to G:C over G:C to A:T transitions could be noted (Table 1).

In conclusion, the GM3 device enables the purging of biofilms to be combined with the programmable delivery of nutrients for an indefinitely long duration.<sup>[7]</sup> Its operation was found to be sufficient for effecting the transliteration of a canonical base with a close chemical analogue over 1000 and 2000 generations, depending on the imposed growth rate. It would have been impossible to predict the genetic alterations underlying these adaptations from current biological knowledge so as to implement them through genome rewriting.

As a replacement for 5-methyluracil (thymine) in DNA, the pairing equivalent found in RNA, uracil, which is unsubstituted at the 5-position, was not adopted, but instead the nonnatural ersatz chlorouracil, which is substituted at the 5-position with a halogen atom that is slightly smaller than the methyl group. In preserving the functional trait of a substituted 5-position on an artificial analogue rather than resorting to a building block devoid of that trait but available in metabolism, we enforced an adaptive process that obeys

the rule of “constraint propagation” commonly encountered in technological evolution.<sup>[31]</sup>

The element chlorine does not occur in the chemical composition of wild-type *E. coli*, yet has become essential for the proliferation of its descendants CLU2, CLU4, and CLU5. Unique molecular interactions mediated by the weak bonds Cl:O, Cl:N, and Cl:S would be disabled by the substitution of chlorine with a methyl group, as in thymine.<sup>[32]</sup> Therefore, it might be expected that further descendants of CLU5 or similar lineages will mobilize and diversify to their advantage the functional usage of chlorine atoms, which has now been imprinted in their genome. Chemically modified organisms, as embodied by our chlorouracil-requiring bacteria, could be systematically diversified in the future to block metabolic cross-feed and genetic cross-talk between synthetic and wild species.<sup>[33]</sup>

In summary, we set out to evolve genomic DNA composed of the three canonical bases adenine, cytosine, and guanine and the artificial base 5-chlorouracil in an *Escherichia coli* strain lacking thymidylate synthase and requiring thymine. Selection over 25 weeks in a cultivation device that automatically adjusts the lowest tolerable thymine concentration yielded descendants that grew with only chlorouracil. The DNA of adapted bacteria contained 90% chlorodeoxyuridine and 10% thymidine. This residual fraction could be forced below 2% by disrupting the *trmA* gene for tRNA U54 methyltransferase, a result that unveiled a cryptic pathway to thymine deoxynucleotides from *S*-adenosylmethionine. Mutations accumulated massively during adaptation to chlorouracil, with a total of 1502 A to G or G to A transitions observed for one culture.

Received: January 21, 2011

Revised: April 15, 2011

Published online: June 27, 2011

**Keywords:** chemical evolution · chlorouracil · DNA · nucleic acids · xenobiology

- [1] A. Kornberg, T. Baker, *DNA Replication*, 3rd ed., W. H. Freeman, New York, **1992**.
- [2] *Metabolism of Nucleotides, Nucleosides and Nucleobases in Microorganisms* (Ed.: A. Munch-Petersen), Academic Press, London, **1983**.
- [3] M. D. Bick, R. L. Davidson, *Proc. Natl. Acad. Sci. USA* **1974**, *71*, 2082–2086.
- [4] E. R. Kaufmann, R. L. Davidson, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 4982–4986.

- [5] R. A. Warren, *Annu. Rev. Microbiol.* **1980**, *34*, 137–158.
- [6] H. H. el-Hajj, L. Wang, B. Weiss, *J. Bacteriol.* **1992**, *174*, 4450–4456.
- [7] P. Marlière, R. Mutzel, Patent WO 00/34433, **2000**.
- [8] a) V. A. de Crécy-Lagard, J. Bellalou, R. Mutzel, P. Marlière, *BMC Biotechnol.* **2001**, *1*, 10; b) P. Marlière, R. Mutzel, unpublished results.
- [9] S. W. Brown, S. G. Oliver, *Eur. J. Appl. Microbiol. Biotechnol.* **1982**, *16*, 119–122.
- [10] P. G. Lane, S. G. Oliver, P. R. Butler, *Biotechnol. Bioeng.* **1999**, *65*, 397–406.
- [11] H. E. Kubitschek, *Introduction to Research with Continuous Cultures*, Prentice-Hall, Englewood Cliffs, NJ, **1970**.
- [12] H. D. Barner, S. S. Cohen, *J. Bacteriol.* **1954**, *68*, 80–88.
- [13] C. E. Deutch, C. Pauling, *J. Bacteriol.* **1971**, *106*, 197–203.
- [14] D. B. Dunn, J. D. Smith, *Biochem. J.* **1957**, *67*, 494–506.
- [15] J. A. Theruvathu, C. H. Kim, D. K. Rogstad, J. W. Neidigh, L. C. Sowers, *Biochemistry* **2009**, *48*, 7539–7546.
- [16] D. W. Visser, D. M. Frisch, B. Huang, *Biochem. Pharmacol.* **1960**, *5*, 157–164.
- [17] L. C. Blasiak, F. H. Vaillancourt, C. T. Walsh, C. L. Drennan, *Nature* **2006**, *440*, 368–371.
- [18] M. Krych, I. Pietrzykowska, J. Szyzko, D. Shugar, *Mol. Gen. Genet.* **1979**, *171*, 135–143.
- [19] F. R. Blattner, G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, Y. Shao *Science* **1997**, *277*, 1453–1462.
- [20] D. J. Porter, B. M. Merrill, S. A. Short, *J. Biol. Chem.* **1995**, *270*, 15551–15556.
- [21] J. Becker, M. Brendel, *Biol. Chem. Hoppe-Seyler* **1996**, *377*, 357–362.
- [22] B. Weiss, *J. Bacteriol.* **2007**, *189*, 2186–2189.
- [23] H. R. Nielsen, K. E. Sjølin, K. Nyholm, B. S. Baliga, R. Wong, E. Borek, *Cancer Res.* **1974**, *34*, 1381–1384.
- [24] G. R. Björk, F. C. Neidhardt, *J. Bacteriol.* **1975**, *124*, 99–111.
- [25] K. A. Datsenko, B. L. Wanner, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6640–6645.
- [26] C. T. Madsen, J. Mengel-Jorgensen, F. Kirpekar, S. Douthwaite, *Nucleic Acids Res.* **2003**, *31*, 4738–4746.
- [27] Y. Motorin, F. Lyko, M. Helm, *Nucleic Acids Res.* **2010**, *38*, 1415–1430.
- [28] A. S. Bhagwat, A. Sohail, R. J. Roberts, *J. Bacteriol.* **1986**, *166*, 751–755.
- [29] C. H. Kim, A. Darwanto, J. A. Theruvathu, J. L. Herring, L. C. Sowers, *Chem. Res. Toxicol.* **2010**, *23*, 740–748.
- [30] M. Hanus, M. Kabelác, D. Nachtigallová, P. Hobza, *Biochemistry* **2005**, *44*, 1701–1707.
- [31] J. P. Tsang, *Lecture Notes in Computer Science* **1990**, *462*, 135–151.
- [32] P. Auffinger, F. A. Hays, E. Westhof, P. S. Ho, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16789–16794.
- [33] P. Marlière, *Syst. Synth. Biol.* **2009**, *3*, 77–84.