Increasing the genetic stability of reduced genome *Escherichia coli* by elimination of the error-prone DNA polymerases

Ph. D. thesis summary

Bálint Csörgő

Supervisor: György Pósfai, Ph. D.

Biological Research Center of the Hungarian Academy of Sciences

Ph. D. School of Biology
University of Szeged, Faculty of Science and Informatics

2012
Szeged

Introduction

The focus of our group for the past several years has been the precisely planned, gradual reduction of the genome of the bacteria *Escherichia coli*. Today, the genomes of over one hundred different strains of *E. coli* have been sequenced. Comparison of the genomes revealed that the genome consists of a set of genes common to all of the strains (the so-called "core" genome) and a set of strain-specific genomic islands acquired mainly by horizontal gene transfer (the so-called "peripheral" genome). Our hypothesis was that the elimination of this latter set of strain-specific genes, mainly needed for the niche-specific adaptation of the different strains, was dispensable under stable, laboratory conditions. In fact, the elimination of these peripheral genomic segments could "streamline" the cells in regards to their metabolic simplicity and efficiency.

We have determined through our work that the elimination of these peripheral genomic elements significantly stabilizes the genome of *E. coli* under laboratory conditions. Molecular mechanisms generating genetic diversity are the foundations for evolution under changing conditions, these provide the basis for the survival of the bacterial population. However, under stable, laboratory conditions, the ability to generate large genetic diversity is not a necessity. In fact, novel, evolved features arising in a carefully designed and fabricated system of biological parts can lead to unwanted genotypic and phenotypic alterations, and the spontaneous genetic modification of an established production strain or a clone library is usually highly undesirable. Consequently, whether used as a production strain, a cloning host, or as a synthetic biological chassis, a bacterial cell with increased genetic stability is of great importance. With this goal in mind, the reduced-genome *E. coli* strain MDS42 (multiple-deletion series strain carrying 42 deletions compared to the wild-type K-12 MG1655 parental strain) was constructed in our laboratories. Most genes irrelevant for laboratory applications, as well as all known mobile DNA sequences and cryptic virulence genes were precisely deleted, resulting in a genetically stabilized strain that displays several advantageous properties. By deleting all mobile genetic elements, the entire insertion class of mutations could be eliminated from the cells. However, point mutations represent a much more common source of genetic diversity in bacterial cells. Unlike insertional

events, all point mutations within the cell cannot entirely be eliminated. Nonetheless, any reduction in the rate of point mutations would still be advantageous for the various applications of the cell.

Our strategy for reducing point mutation rate in *E. coli* involved disabling the effective mutation generating enzymes of the SOS response. Under stressful conditions (e.g., toxic clones harbored in the cell), DNA damage may occur, activating the SOS response, inducing approximately 40 members of the SOS regulon. Three of the genes induced during the SOS response of *E. coli* encode DNA polymerases (Pol II, Pol IV, Pol V) that are able to bypass replication barriers at damaged sites and stalled replication forks. All three of the SOS-inducible polymerases have been implicated in induced mutagenesis, with Pol IV and Pol V having error-rates approximately 2 to 3 orders of magnitude higher than the high-fidelity replicative polymerase (Pol III). Pol II, while showing high fidelity on undamaged templates, was shown to take part in certain types of stress-induced mutagenesis. The SOS-regulated, pointmutation-generating polymerases are dispensable; their primary role seems to be the generation of genetic diversity under stressful conditions. However, it was unclear whether elimination of these polymerases, components of the DNA repair mechanisms, results in growth defects, particularly under stress.

Aims

The primary goal of our work was the development of a genetically highly stable *Escherichia coli* strain with a reduced mutation rate. By eliminating different variation generating mechanisms of the parental strain, we hoped to produce a strain that could be advantageous in both laboratory and industrial applications. Besides these applied goals, basic questions regarding the mutation generating mechanisms of the cell could be answered.

By eliminating the genes coding for the SOS-inducible error-prone DNA polymerases of *E. coli*, we wished to address the following specific questions:

- To what extent do the error-prone polymerases of E. coli play a role in the generation of spontaneous mutations?
- Does the mutation rate of the cell increase under various stress conditions and what effect do the error-prone DNA polymerases have on this?
- Does inactivating the SOS response have the same effect as the deletion of the genes coding for the error-prone DNA polymerases?
- How does the absence of error-prone DNA polymerases affect the survival properties of the cell?
- Does the absence of the error-prone DNA polymerases help in the stable propagation of a toxic clone within the cell?

Methods

- Recombinant-DNA techniques for the synthesis of plasmids used for genome modification
- Transformation of bacterial cells
- Bacterial growth rate measurements
- Mutation rate measurements
- Mutational spectrum analysis
- Overexpression of proteins in bacteria
- DNA sequence analysis
- Statistical testing

Results

The genes coding for the three inducible DNA polymerases (*polB, dinB, umuDC*) were deleted from the genome of MDS42 in a scarless manner using a suicide plasmid-based method. Gene deletions were made individually and also in all possible combinations. The deletion of each error-prone polymerase gene by itself resulted in at least a 20 % decrease in spontaneous mutation rate. When combining the different deletions, the

mutation rate decreased even further, the spontaneous mutation rate of the triple deletion strain MDS42*polBdinBumuDC* (MDS42pdu) was approximately 50 % lower than that of the parental MDS42 strain. It is important to note that the basic fitness properties of the MDS42pdu strain, such as growth rate and survival during long-term stationary phase were not significantly different from those of its parental MDS42 strain.

Due to the stress-induced nature of the error-prone DNA polymerases, it was expected that the difference in mutation rates of the polymerase-free and the parent strain would be even more pronounced under stressful conditions. A sub-inhibitory concentration of the SOS response-activating mitomycin-C, overproduction of either the non-toxic GFP protein or of the highly toxic ORF238 hydrophobic protein all significantly increased the mutation rate of MDS42. The values for MDS42pdu remained stable under the same conditions. It is also noteworthy, that among the strains tested, the commonly used production strain BL21(DE3) showed not just the highest spontaneous mutation rate, but also the highest increase in mutation rates in response to the various stresses. A difference of almost two orders of magnitude was observed between the mutation rate of BL21(DE3) and MDS42pdu when overproducing the toxic ORF238 protein. This elevated rate of mutation in BL21(DE3) can be mostly attributed to an increased rate of IS insertions.

A clear practical advantage of working with MDS42pdu was demonstrated in a protein production experiment, where the SinI methyltransferase was expressed from an inducible plasmid construct. M.SinI, producing 5-methylcytosines is toxic to cells that carry the McrBC endonuclease. Even in cells lacking McrBC, we observed a negative effect on cell fitness. When M.SinI was produced, we found that the *sinIM* gene, carried on a plasmid, acquired loss-of-function mutations approximately three times less frequently in MDS42pdu than in MDS42, and over five times less frequently than in BL21(DE3)*mcrBC*. Remarkably, after only 16 hours of production in BL21(DE3)*mcrBC*, almost half of all *sinIM* genes encoded on the plasmids had suffered a disabling mutation.

Clearly, the unexpectedly high ratio of mutated clones in the M.SinI-expressing culture cannot be explained solely by the stress-induced mutagenesis, the overall mutation rate of which being too low in absolute values (in the order of 10^{-6} mutations/gene/generation) to cause such a dramatic effect. Rather, the phenomenon is in

large part due to the growth inhibitory effect of the plasmid carrying the toxic gene. The chain of events could be the following: Upon induction of expression of the toxic gene, growth rate of the cell is reduced. At the same time, mutation rate is increased by the stress. Once a mutant, not producing the toxic protein, arises in the plasmid population, the cell harboring it can resume normal growth and become dominant in the culture. In low-mutation-rate MDS42pdu, appearance of such mutants is delayed, and the cells can produce the functional toxic protein for an extended period of time.

The mutation and inactivation of engineered genetic constructs within a host cell is an overlooked problem that may have serious detrimental effects on the success of any synthetic biological, molecular biological or biotechnological process. A gene product imposing a metabolic burden or being toxic to the host drives an evolutionary force that selects for any mutants that alleviate the growth-inhibiting effect. A host cell or chassis with enhanced genetic stability is advantageous in the stable maintenance of these constructs. By eliminating the inducible DNA polymerase enzymes from the reduced-genome MDS42 strain lacking all genomic IS elements, we have further stabilized a strain that already showed clear advantages in cloning applications. The resulting MDS42pdu strain had a significant stabilizing effect on a toxic protein expression clone. This high-fidelity strain, producing decreased genetic variation in the culture, might also prove useful in applications ranging from the production of DNA therapeutics to long-term continuous fermentation processes.

Results in brief

We have generated variants of a reduced genome *Escherichia coli* strain constructed previously in our laboratory lacking the SOS-inducible error-prone DNA polymerase enzymes (Pol II, PolIV, and Pol V).

We show that both the spontaneous and stress-induced mutation rate of the strain lacking all error-prone DNA polymerase enzymes (MDS42pdu) is lower than the parental strain and its *recA* and *lexA* mutant variants.

We show that the error-prone DNA polymerase-free strain allows for the propagation of a toxic protein-coding plasmid with increased stability.

Our results show that our MDS42pdu strain of *E. coli* is a better, more stable cloning host than those previously used.

List of publications

1. Csörgő B, Pósfai G: **Directed homologous recombination for genome engineering in Escherichia coli**. *Acta Biol Hung* 2007, **58 Suppl**:1-10.
   I.F.: 0.447
2. Fehér T, Karcagi I, Győrfy Z, Umenhoffer K, Csörgő B, Pósfai G: **Scarless engineering of the Escherichia coli genome**. *Methods Mol Biol* 2008, **416**:251-259.
3. Durfee T, Nelson R, Baldwin S, Plunkett G, 3rd, Burland V, Mau B, Petrosino JF, Qin X, Muzny DM, Ayele M, Gibbs RA, Csörgő B, Pósfai G, Weinstock GM, Blattner FR: **The complete genome sequence of Escherichia coli DH10B: insights into the biology of a laboratory workhorse**. *J Bacteriol* 2008, **190**(7):2597-2606.
   I.F.: 4.013
4. Csörgő B, Fehér T, Tímár E, Blattner FR, Pósfai G: **Low-mutation-rate, reduced-genome Escherichia coli: an improved host for faithful maintenance of engineered genetic constructs**. *Microb Cell Fact* 2012, **11**:11.
   I.F.: 4.544
5. Fehér T, Bogos B, Méhi O, Fekete G, Csörgő B, Kovács K, Pósfai G, Papp B, Hurst LD, Pál C: **Competition between transposable elements and mutator genes in bacteria**. *Mol Biol Evol* 2012, In press.
   I.F.: 8.907