The Genetic Basis of the Correlation Between DNA Replication Regulation, Central Carbon Metabolism and Stress Alarmons in *Escherichia coli* Cells

Monika Maciąg-Dorszyńska

To survive, cells have developed complex systems to adjust and coordinate their main biological activities to the energy provided. One of the cellular processes which precise regulation and coordination is fundamental to life is DNA replication. This process leads to duplication of genetic information prior to cell division. Duplication of the genetic material of bacteria starts in the region named *origin* (for *Escherichia coli* - *oriC*) and extends bidirectional up to terminatory *ter* region. Completion of correct process of DNA replication requires a number of proteins that form complex biological structure. This structure consists of the DnaA initiator protein, DnaB helicase, helicase transporter DnaC, DNA polymerase III holoenzyme (formed by at least 10 subunits), primase DnaG, and gyrase. Moreover, there are several other DNA polymerases in *E. coli*: polymerase I of 5'->3' polymerization, as well as 5'-> 3' and 3'-> 5' DNA exonuclease activities, which main functions is removal of primers, joining Okazaki fragments and DNA repair, DNA polymerase II which has 3'->5 exonuclease activity and participates in DNA repair, DNA polymerase IV, engaged in non-targeted mutagenesis, DNA polymerase V, involved in the SOS response and translesion synthesis during DNA repair.

DNA replication consists of three stages: initiation, elongation and termination. Within a scope of regulatory control, initiation and termination stages are the most known steps of DNA replication. Regulation of elongation is not well known although this step is a key element in the cell function. Relations between nutrients availability affecting cell metabolism and DNA replication could be a key element in maintaining genome stability. DNA replication as well as the cell cycle depends on nutrients' availability. Interestingly, the cell cycle of *E. coli* and some others bacteria in poor media is similar to eukaryotic cell cycle. Cell cycle of slow growing bacteria consist of three periods: (i) the time from cell division to initiation of chromosome replication, (ii) the time for chromosome replication and (iii) time from termination of replication to cell division. In rich media, a new round of replication starts before previous one is completed which results in reduction of cell cycle duration. The differences between cell cycles in different media suggest a connection between DNA replication and cell metabolism.

The core of the cellular metabolism is the central carbon metabolism (CCM). CCM is a series of processes that converts nutrients from environment to power cell metabolism. Intermediate metabolites produced in these processes are used for synthesis of macromolecules. Glycolysis/gluconeogenesis, pentose phosphate pathway and Krebs cycle constitute the core of central carbon metabolism, while glucose is the most common substrate used in high energy production. Nevertheless, the cell can use a variety of organic compounds and incorporate them to the advantageous stage of pathways. Dysfunction in central carbon metabolism pathways may result in cell cycle and DNA replication deregulation. Depending on stage at which this disruption occurs, cell could not divide, replicate their genetic material or initiate all mechanisms responsible for fidelity of the replication process. However, the mechanism connecting DNA replication and the central carbon metabolism remains largely unknown.

We can speculate that this connection may include undiscovered enzymatic function, and this hypothesis can be supported by the observation that some of metabolic proteins can form non-metabolic structures and functional complexes. CCM enzymes may have non-metabolic, catalytic activity (uracil glycosidase, protein kinase) associated with DNA repair. Moreover, bacterial virulence depends on glycolytic enzymes located on cell surface. In many diseases, these enzymes can interact with proteases to degrade mucous membrane for pathogen's invasion. Link between central carbon metabolism and DNA replication can also relate to mechanisms which are not yet identified and may influence the cell cycle and DNA replication.

The first aim of my work was identification of potential relationships correlating central carbon metabolism (glycolysis/ gluconeogenesis, Krebs cycle, penthose phosphate pathway and pyruvate metabolism) with DNA replication in *Escherichia coli*. A study conducted in 2007 by Dr. Janniere's team identified genes encoding CCM enzymes which can modulate the activity of the replication apparatus in *Bacillus subtilis* bacterium. Notably, some mutations in genes coding for enzymes involved in glycolysis suppressed replication defects. This suggests that some metabolic enzymes can modulate replisome properties in response to physiological state of the cells. Thus I asked if this is a specific phenomenon, characteristic for this bacterium and perhaps for its close relatives, or a more general biological rule.

To find an answer for this biological question I constructed *E. coli* strains with combination of mutations in replication genes (conferring temperature-sensitive phenotype) and in genes encoding metabolic enzymes. For this, I employed bacterial strains with

mutations in genes encoding DnaA (a protein responsible for the initiation of replication), DnaC (DnaB helicase transporter protein), DnaG (primase), DnaE (alpha subunit of DNA polymerase III) and DnaN (beta subunit of DNA polymerase III) and the set of strains with deletion mutations in genes coding for specific metabolic enzymes obtained from Keio collection. I analyzed suppression of temperature-sensitive effects of mutations in genes encoding DNA replication proteins by mutations in genes involved in central carbon metabolism pathways (glycolysis/gluconeogenesis, pentose phosphate pathway, pyruvate metabolism and the Krebs cycle). The effect of *dnaA46*(ts), *dnaG*(ts), *dnaN156*(ts) mutations could be suppressed by dysfunction of *pta* or *ackA* from pyruvate metabolism pathway. Effect of *dnaB8*(ts) mutation was suppressed also by dysfunction of *pta* and *pgi* which products catalyze gluconeogenesis and glycolysis reactions, effect of *dnaE486*(ts) by dysfunction of *tkt*B, effect of *dnaG*(ts) also by dysfunction of *gpmA* from pentose phosphate pathway and glycolysis/gluconeogenesis pathways respectively [1].

The process of enzymatic DNA synthesis is error-prone, which results in the appearance of mutations. To avoid spontaneous mutations, bacteria evolved specific mechanisms which are responsible for fidelity of DNA replication. High accuracy of DNA replication is primarily maintained by the presence of alpha (DnaE) and epsilon (DnaQ) subunits of DNA polymerase III. DnaE is a catalytic subunit, which catalyses the DNA synthesis reaction, but at the same time providing fidelity of nucleotide insertion, while DnaQ is a 3'→5' exonuclease with the proofreading activity. One of the commonly used mutants in the gene coding for the epsilon subunit, *dnaQ49,* is a recessive allele that confers an impaired proofreading phenotype. It is characterized by a high rate of spontaneous mutations and chronic induction of the SOS response. In addition to *dnaE* and *dnaQ*, *dnaX* is another gene coding for DNA polymerase III subunit whose products take place in DNA replication fidelity control. Recent studies indicated that this fidelity function applies to both leading- and lagging-strand synthesis.

Commonly used mutants in the genes coding for the epsilon and tau subunits are *dnaQ49* and *dnaX36*, respectively. The presence of each of this alleles leads to appearance of a high number of spontaneous mutations. In this part of my work I asked whether the link between CCM and DNA replication is restricted to efficiency of synthesis of polynucleotide strands or it can be also applied to another feature of this process, namely fidelity of the replication. Determination of frequency of mutations showed that the increase in the mutator effect in strains harboring *dnaQ49* and *dnaX36* alleles were suppressed by the presence of mutations in the genes coding for aconitate hydratase 2 (AcnB), isocitrate dehydrogenase

specific for NADP (Icd), acetate kinase (AckA) and phosphate acetyltransferase (Pta). Dysfunction in the *dnaX* gene was suppressed also by the mutation in gene encoding glucose-6-phosphate-1-dehydrogenase (Zwf). The effects of these suppressions could be reversed by plasmid-mediated expression of the wild type CCM genes [2].

Changes in environmental conditions constitute to grow rate dysfunction. This effect could be connected with changes in the cell cycle, cell mass doubling, chromosome replication and chromosome segregation in bacterial cell. The replication deregulation can result in severe defects in e.g. cell morphology (if cells do not divide correctly and make filaments) or chromosomes' segregation during cell division. Thus, the next part of my work was aimed at revealing the mechanism of the replication and carbon metabolism link, and involved analysis of *dna*Ts and *dna*Ts suppressor mutants' parameters. For this, I performed microscopy-based analysis of cell morphology, and nucleoid shape and position [3].

Length of the *dnaA46*, *dnaB8*, *dnaG* (ts), and *dnaN159*, *dnaE486* mutants' cells varies from 10 to 30 μm. The presence of the additional mutations in *pta*, *ackA*, *pgi*, *tktB* and *gpmA* genes, suppressed filamentation effect and nucleoid disruption. Similarly to the previous observations, this suppression could be reversed by plasmid-mediated expression of the wild type CCM genes.

Many changes in environmental conditions could prompt to nutrient deficiency. This effect can also activate specific adaptive mechanisms. One of them is the stringent response, whose effectors, specific nucleotides, guanosine tetraphosphate and guanosine pentaphosphate (p)ppGpp) are rapidly produced in response to a variety of physico-chemical and nutritional stresses. These nucleotides can regulate directly or indirectly various processes such as gene expression (by interaction with RNA polymerase), translation and initiation of DNA replication. The second transcriptional factor, taking part in stringent response is DksA protein. DksA enhances the effect of guanosine tetraphosphate on ribosomal promoters and positive regulates the amino acids synthesis in cooperation with ppGpp. DksA together with ppGpp may also take part in DNA repair. The results presented by Dr. Wang and coworkers showed that ppGpp can regulate replication elongation in *B. subtilis* by inhibiting primase, which activity is necessary for RNA starters' synthesis on leading and lagging stands. Since no stringent control-mediated replication forks arrest in *E. coli* was reported, an interesting question appeared whether (p)ppGpp-mediated inhibition of DnaG primase activity is restricted to *Bacillus subtilis* or it is a more general phenomenon and can occur also in other bacteria.

In vitro primer synthesis analysis showed that guanosine tetraphosphate (ppGpp) significantly reduces the DnaG activity. Less significant inhibition was observed in the presence of guanosine pentaphosphate (pppGpp). These effects were observed both in presence and absence of DnaB helicase [4]. This suggests a direct effect of ppGpp on DnaG primase. Since DksA is considered as a co-factor of the stringent response, we asked whether it can affect DNA replication [5]. The addition of purified DksA to the *in vitro* replication assay showed no effect of this protein. Analysis of the kinetics of DNA replication *in vivo* using ppGpp-deficient strain, under conditions of amino acid starvation, showed no significant difference in DNA replication in relation to the wild-type strain. These results may be an effect of many processes. One possibility is that *in vivo* there is/are factor(s) masking effects of ppGpp on DnaG primase by preventing its binding to this protein. Such factor(s) would be absent in our *in vitro* assays for measurements of kinetics of primase activity and DNA synthesis. Another hypothesis which may explain the mechanism of masking the ppGpp potential to impair DnaG activity is that the primase is not efficiently inhibited in *E .coli* cells due to competition for binding of ppGpp to this protein and to RNA polymerase.

Results presented in this work showed correlation between crucial cellular processes. However, the underlying mechanism remains yet unknown. Results presented in my PhD thesis constitute a basis for the continuation of research on mechanisms connecting DNA replication and central carbon metabolism.

References

[1] Maciąg M., Nowicki D., Janniere L., Szalewska-Pałasz A., Węgrzyn G. 2011. Genetic response to metabolic fluctuations: correlation between central carbon metabolism and DNA replication in *Escherichia coli*. Microbial Cell Factories 10: 19.

[2] Maciąg M., Nowicki D., Szalewska-Pałasz A., Węgrzyn G. 2012. Central carbon metabolism influences fidelity of DNA replication in *Escherichia coli*. Mutation Research 731: 99-106.

[3] Maciąg-Dorszyńska M., Ignatowska M., Jannière L., Węgrzyn G., Szalewska-Pałasz A. 2012. Mutations in central carbon metabolism genes suppress defects in nucleoid position and cell division of replication mutants in *Escherichia coli*. Gene 503: 31-35.

[4] Maciąg M., Kochanowska M., Łyżeń R., Węgrzyn G., Szalewska-Pałasz A. 2010. ppGpp inhibits the activity of *Escherichia coli* DnaG primase. Plasmid 63: 61-77.

[5] Maciąg-Dorszyńska M., Szalewska-Pałasz A., Węgrzyn G. 2013. Different effects of ppGpp on *Escherichia coli* DNA replication *in vivo* and *in vitro*. FEBS Open Bio 3: 161-164.