IS CELL CYCLE A PERFECT FAULT TOLERANT CONTROL SYSTEM?
– SYSTEM ENGINEERING POINT OF VIEW

Andrzej Świerniak, Roman Jaksik, Jaroslaw Śmieja
Department of Automatic Control, Silesian University of Technology
Akademicka 2A, 44-100 Gliwice, Poland
e-mail: andrzej.swierniak@polsl.pl, roman.jaksik@polsl.pl, Jaroslaw.Smieja@polsl.pl

ABSTRACT
We present a system engineering approach to a problem of DNA damage-repair in human cells. We discuss chosen control systems existing in the cells and demonstrate that they constitute an almost perfect fault tolerant system. We discuss the consequences of gaps in this system and explain why it cannot be completely perfect.

KEY WORDS
Cell cycle, DNA damage-repair, fault-prone system, p53 regulatory module

1. Introduction
Mathematical modeling and use of advanced information processing methods have become an integral part of the molecular biology research in recent years. Due to complexity of regulatory networks, governing all intracellular processes, it is not possible to rely on experimental work only, which does not utilize one of the many mathematical tools available (systems modeling and analysis, graph theory, statistical inference, etc.) to expand knowledge of these processes.

This paper is concerned with cell cycle regulation, as one of the most important of these networks and presents a unique point of view, looking at it as an engineering construct.

2. Intracellular information processing
Human cell is a very complex biological system responsible for the appropriate functioning of its internal processes and for maintaining the cellular population in which it resides. Each cell consists of cytoplasm and organelles which are used to conduct various metabolic processes required for the cell and organism survival, its ability to develop and reproduce. The central element of the cell is the nucleus which stores the genetic information used to build individual elements of the cell structure and control all of the chemical and biological processes. The information contained in the nucleus is stored using deoxyribonucleic acid (DNA) based on four distinct subunits termed nucleotides. The specific order of nucleotides additionally controls the information availability and rate of transcription, a process in which its specific elements termed genes are copied to a template build from ribonucleic acid (RNA). This template is used to create multiple copies of a single protein in a process of translation. An essential feature of this mechanism is that its efficiency depends not only on the complexity of the RNA template and availability of the gene region but on the concentration and activity of various other molecules involved in the production process [1]. Additionally each template can be used multiple times until it degrades and multiple various templates can originate from a single gene.

This results in a very complex regulation system that allows the cell to conduct various processes based on its current requirements, allowing to maintain the cellular structure [1], catalyze biochemical reactions [2], carry molecules from one place to another [3], control the efficiency of gene expression [4] and control the stability of other molecules [5]. The control system is able to respond to various external signals using a complex mechanism of signal propagation which is based on the balance between production and degradation rate of various molecules (Fig.1). The basic mechanism of the production rate control utilizes a set of specific proteins termed transcription factors (TF) that initiate the RNA production process by binding to specific regions in the DNA. Their amount, specificity to the DNA and accessibility of binding regions determine the rate of transcription. The negative control is much more complex, occurring on various levels of the gene expression process. Negative regulation is based on control of the mRNA decay, using either RNA binding proteins (RBP) or short non-coding RNAs.

All of the regulatory mechanisms introduced are based on a target recognition processes that requires a specific nucleotide sequence motif in order for the regulation to occur. The required specificity level varies between protein- or RNA-based regulatory factors and depends on a formation of many week non-covalent bonds, including ionic bonds, hydrogen bonds, and van der Waals attractions.
Human genome includes over 22 thousand genes which can be used to create over 45 thousand distinct RNAs used as a template for the creation of over 500 thousand distinct proteins, that can be further modified post-translationally. Post-transnational protein modifications form a next class of the signal propagation system. Chemical modification of the protein can affect its activity, cellular location and stability, which allows to control various intracellular processes. Such processes include cell differentiation, metabolism, or immunological response. Phosphorylation and ubiquitination are two most common protein modifications which can turn on or off specific functions of various proteins by changing their activity or affecting molecule binding capabilities [6].

Protein modifications extend the regulatory capabilities of the cell, creating an incredibly sophisticated system. The control over individual biological processes and levels of particular proteins is maintained through many positive and negative feedback loops build of various RNAs and protein molecules that interact with each other. Negative feedback loops are an essential element of intracellular regulation systems in all complex organisms, allowing to maintain homeostasis by controlling the levels of synthesis and turnover rate of various chemical substances. Positive feedback loops are not as common although they are still an essential element of the regulatory system. Cellular positive feedback loops work by either activating protein A trough protein B which is responsible for protein A activation or more commonly trough double negation where A inhibits an A-inhibiting factor B [7]. Positive feedback is used to invoke a very fast response for example if DNA damages were detected [8, 9]. Such mechanism is utilized in many signaling pathways and plays an important role in the regulatory processes of the cell cycle, which allows the cell to grow and divide making it one of the most complex regulatory systems found in nature.

3. Control of the cell-cycle

Cell cycle is a sequential process which due to its time-dependency and repeatability is often compared to a clock (see [10] and references therein). The entire process is controlled by specific checkpoints which stop the cycle unless a specific “go” signal is received. This allows to control timing of each stage and their order, additionally making sure that the transition to the next stage will not occur if the previous one was not completed successfully. Cell cycle time varies significantly depending on cell type, although for most of the human cells it ranges from 18 to 24 hours and can be divided into four distinguishable phases G1, S, G2 and M. During the first phase (G1) cell increases its size and produces proteins required in the synthesis (S) phase during which DNA is replicated in order to pass a full set of chromosomes to the daughter cells. G2 is the second growth phase that prepares the cell for the mitosis (M) phase during which the cell divides separating the newly replicated genetic material (Fig. 2).

4. Cell cycle checkpoints

The robustness of the cell cycle is ensured thought a variety of backup mechanisms, control switches and checkpoints that respond to abnormal cell conditions or environmental stress. The cycle includes a total of three checkpoints, located at the ends of G1, G2 and M-phases (Fig. 2), which...
ensure that the next process can be initiated only if the previous one was completed successfully and that the cell condition is suitable to carry on to the next step [11]. G1 checkpoint (checkpoint R) prevents DNA replication if its structure is damaged. The damage level is controlled thought various proteins that detect single or double strand breaks and various chemical modifications like depurinations or cytosine deamination [12]. The damage detection proteins activate a specific protein kinase – ATM [9] which controls the cell cycle arrest mechanism, providing time necessary to initiate the DNA repair processes [13]. If the damages are beyond repair or the repair process exceeds the provided time the cell will switch to a non-dividing state termed G0. G1 checkpoint prevents premature transition to the next replication step and propagation of DNA damages to the daughter cells [13]. G2 checkpoint has a similar role. By monitoring the DNA integrity during the G2 phase it ensures that the DNA is properly replicated and that the replication completed, before entering mitosis phase. It does that by inhibiting B1/Cdc2 cyclin which prevents the cell from entering M-phase until the replication is complete and/or the DNA damages are fixed [14]. The final M-phase checkpoint controls the genome integrity after mitosis ensuring that each daughter cell received a complete copy of newly replicated chromosome [15]. Cell cycle checkpoints prevent genomic instability which occurs when the daughter cells receive only a part of the DNA or the DNA is severely damaged. The mechanism which controls this process is highly responsive to information obtained from each of the control steps, dynamically changing the course of the entire process based on the current state of the cell. The fault resistance of the system is achieved by the use of multiple negative signals that block the advancement to the next stage rather than positive, that stimulate the cycle progression. Each of the negative signals is activated independently if at least one of the conditions is not met, for example the level of required proteins or nutrients is too low for the cell to conduct the process or the growth of the cell and its subcomponents was not completed [1]. From the control theory point of view such approach has a significant advantage since the system is much more fault resistant if it has to detect at least one “stop” signal rather than multiple “go” signals indicating that the previous processes were completed successfully. The effectiveness of this mechanism depends however on the ability of the cell to quickly detect damages to its internal components and most importantly on the ability to control the genome integrity in which p53 protein signaling pathway plays a significant role.

5. Damage recovery

Despite the redundancy in multiple mechanisms described above, the cellular detection and repair system is not failure-free. As a result, mutations, rearrangements and other disruption of genetic information may take place, particularly during DNA replication. They are the most dangerous when they occur in genes that code for proteins controlling either cell cycle or DNA repair processes, since in these cases the error signal propagates through the cell gene regulatory network an later may spread in a population, leading to development of cancer. In the case of genes involved in repair processes, the result is loss of genomic stability. Two critical elements are necessary in the control mechanisms described in the preceding section. The first one is the subsystem used for fault detection. Though base pair mismatch and single strand breaks might be sensed through several pathways, double strand break recognition relies on ATM molecules. This mechanism is broken in several diseases, including leukemia and lymphoma.

Another sensitive component can be found in repair subsystem. Mutations in mismatch repair genes are involved in development of diffuse large B cell lymphomas, colorectal tumors, some forms of breast, ovarian and pancreatic cancer are associated with mutations in recombination modifying genes. The negative effects of inefficient or improperly working failure detection and repair subsystems can bring even more damage if combined with malfunctioning p53 pathway. Actions exerted by the tumor suppressor p53 should counteract the accumulation of DNA defects leading to cancer. Therefore, if the p53 protein is dysfunctional, e.g. due to mutation in the gene that codes it, cells carrying wrong genetic information may divide give rise to various forms of human cancer. Damages to the genetic material occur very often, for example Kohn and Bohr [16] report around 55 thousand single strand breaks per cell per day. For this reason cell cycle must be able to withstand damages that can occur during each of the cell division stages and prevent them from being propagated, which could have catastrophic consequence for the organism. DNA damages occur most commonly as a result of exposition to genotoxic substances and factors such as ultraviolet [17] and ionizing radiation [18] or high level of reactive oxygen species [19]. Damage detection is not trivial due to the size of the DNA which in human cells exceeds 3 billion base pairs, requiring a very sensitive detection mechanism, which could not only detect single- or double-strand breaks but also modifications of certain DNA monomers [14]. DNA damages trigger the p53 protein signaling pathway, which controls various processes that protect the cell and if necessary activate its self-degradation [8]. It does that by initiating cell cycle arrest in G1 or G2 phase and triggering the DNA repair machinery if severe damages were detected [8]. One should also remember that environment also contributes to stability of the cell cycle [2].

Extracellular signals from neighboring cells, called mitogens may overcome intracellular mechanisms that block or slow the cell cycle. They act through signaling pathways involving a small GTPase Ras. A mutation in Ras-coding gene [20] may cause it to be permanently active, thus leading to continuous progress of the cell cycle as well as helping to fuel metabolic pathways, supporting growth and division. These mutations are found in about
25% of human cancers and are highly prevalent in hematopoietic malignancies. On the other hand, viral infections may also lead to changes that promote activation of transition to the next phases of the cell cycle under wrong conditions, and, ultimately, result in carcinogenesis. For example, human papilloma virus (HPV) produces oncoproteins E6 and E7, which disrupt, otherwise well-performing, regulatory network. The first of these blocks p-53 mediated activation of the p21 protein, while the other inactivates Rb, thus activating E2F and inducing cell cycle progression independent of the G1-S checkpoint Cdk s.

Fig. 3. Simplified model of the p53 protein regulation

The p53 pathway regulation involves two feedback loops one positive and one negative [21], as shown on Fig. 3. The first, negative loop involves Mdm2 protein which is responsible for p53 degradation [22]. Mdm2 is however activated by the p53 creating a feedback which maintains a constant level of p53 in the cell. The second, positive loop, involves AKT protein and works thought double negation. AKT is negatively regulated by the p53 and at the same time it mediates the Mdm2 dependent p53 degradation. This mechanism allows p53 to inhibit its own inhibitor, which, if the DNA damages are detected and p53 is triggered by the ATM, significantly increases the p53 concentration [23]. ATM additionally can block the Mdm2-mediated p53 degradation, which allows to significantly increase the p53 level [9]. If activated p53 pathway provides the cell only a certain time to repair the DNA damages. If the repair is ineffective p53 level reaches a very high level triggering the transcription of proapoptotic genes that initiate the programmed cell decay named apoptosis. This mechanism is used to eliminate severely damaged cells, protecting the entire organism. From system engineering point of view this system is a control system with switching parameters (see e.g. [24]). Analysis of this class of hybrid systems and synthesis of switching rules for them belong to the hottest issues in modern control theory see e.g. [25]).

Replication of damaged DNA is not the only problem that may arise in cell cycle and contribute to the development of cancer. Another one lies in acceleration of the cell cycle, caused by earlier than necessary entry into the S phase (Fig.4).

Fig.4. Interaction network regulating entry into the S phase and sources of its failure in cancer cells: (A) D1 overexpression; (B) p16 inactivation; (C) Rb inactivation; (D) p21 and (E) p27 failures.

This entry is dependent on increased activity of the E2F protein. It can be promoted either by inducing transcription of E2F gene or by inactivation of the retinoblastoma protein (Rb) that acts as a brake on the cell cycle progression. This, in turn, can be achieved through the activation of the G1-Cdk cyclin (cyclin D-Cdk4), preceded by increased cyclin D1 production. The cyclin D1 is frequently overexpressed in a wide range of cancers, sometimes coincident with gene amplification or somatic mutations of the gene coding it. A frequent alternative splicing leads to production of cyclin D1b protein that lacks a specific phosphorylation site required for nuclear export, leading to its accumulation in the nucleus and increased interaction with Rb, and, subsequently, promoting entry into the S phase [1]. While this could be prevented by another control mechanism, based on the p16 protein that blocks the formation of an active D1-cdk4 complex, many cancer cells have either a deleted, inactivated or silenced p16 gene. Moreover, some mutations in the p16 gene promote cancer metastasis. On top of this, in some cancers p16 is overexpressed and despite that, these cancers may have poor prognosis. This suggests that our knowledge of even this, relatively small part of the signaling network, is far from complete and caution is recommended, concerning conclusions drawn from experimental work and mathematical modeling that supports it. Another important regulator of the cell cycle is the p27 protein. Its increased degradation, natural in a normal cell cycle, leads to G1/S-Cdk activation, thus promoting entry into the S phase. It has been found that in some tumors p27 is mutated in a way that reduces its stability. The ultimate result of such mutation is, once again, uncontrolled entry into the S phase and acceleration of the cell cycle. On the other hand the failure prone system cannot be perfect in order to allow some nucleotide substitutions to occur becoming one of the most important elements of the organism evolution. Permanent changes of the nucleotide sequence occurring as a result of unpaired DNA damages, named mutations are crucial for the survival of the entire species.
control mechanisms. Analysis of the dynamical properties, and comparison of simulation and experimental results help to find missing elements of the signaling network and identify kinetic parameters of the processes [26]. This, in turn, helps to plan subsequent biological experiments and enhance the analysis of external simulation effects on the signaling pathways [27]. Moreover, the system engineering point of view enables to simplify those models in some stages of their analysis without wasting important control properties of the modeled processes. Moreover, it opens new perspectives of controlling these processes by external interventions on the molecular level. Although such mechanisms have been already discussed in literature (see [28] and references therein) their understanding is far from being complete.

**Acknowledgment**

This work was partially supported by BK-213/RAu1/2016, t.3 grant.

**References**


