Author's Review of Her Research Achievements and Publications

1. First and Last Names.
Katarzyna Potrykus

2. Obtained diplomas, scientific/artistic titles and degrees – with their name, place and year of their receiving and title of the doctoral dissertation.

Master of Science in biotechnology – Intercollegiate Faculty of Biotechnology, University of Gdańsk- Medical University of Gdańsk, June 1999.

Ph.D. in biology – University of Gdańsk, Faculty of Biology, Geography and Oceanology, November 2003. Title of doctoral dissertation: " Regulation of transcription and replication of bacteriophage λ – the role of guanosine tetraphosphate (ppGpp) in control of promoters' activities".

3. Information about employment at research/art centers

2003- till now	adjunct, University of Gdańsk, Faculty of Biology, Department of Molecular Biology
(May 2004-May 2012	postdoctoral fellowship at the National Institutes of Health, Bethesda, USA)

4. Description of the 'scientific achievement', as described in Art. 16. Section 2 of the bill published on March 14, 2003, regarding scientific degrees and titles and regarding degrees and titles in the arts (*Dz. U. nr 65, poz. 595 ze zm.*):

a) title of the scientific achievement

The effect of (p)ppGpp and GreA, GreB, and DksA transcriptional factors on RNA polymerase and global cellular regulation in *Escherichia coli*.

b) (author/authors, title of the publication, date of publication, name of publisher)

[1] **Potrykus K**, Vinella D, Murphy H, Szalewska-Palasz A, D'Ari R, Cashel M. (2006) Antagonistic regulation of *Escherichia coli* ribosomal RNA *rrnB P1* promoter activity by GreA and DksA. *J Biol Chem.* 281(22):15238-48. (IF 2006 = 5.808)

[2] **Potrykus K**, Cashel M. (2008) (p)ppGpp: still magical? *Annu Rev Microbiol.* 62:35-51. (IF 2008 = 10.902)

[3] **Potrykus K**, Murphy H, Chen X, Epstein JA, Cashel M. (2010) Imprecise transcription termination within *Escherichia coli greA* leader gives rise to an array of short transcripts, GraL. *Nucleic Acids Res.* 38(5):1636-51. (IF 2010 = 7.836)

[4] **Potrykus K**, Murphy H, Philippe N, Cashel M. (2011) ppGpp is the major source of growth rate control in *E. coli*. *Environ Microbiol.* 13(3):563-75. (IF 2010 = 5.537)

[5] Vinella D, **Potrykus K**, Murphy H, Cashel M. (2012) Effects on growth by changes of the balance between GreA, GreB, and DksA suggest mutual competition and functional redundancy in *Escherichia coli*. *J Bacteriol.* 194(2):261-73. (IF 2010 = 3.726)

c) Aim of the above publications and description of obtained results, including description of their possible utilization.

The aim of the above publications was to expand our knowledge about the effect of bacterial transcriptional factors on RNA polymerase in the light of the stringent and relaxed responses in *Escherichia coli*. Stringent response is an event taking place in cells undergoing different kinds of stresses or starvations (for e.g. starvation for amino acids, carbon, nitrogen, iron, phosphate, lipids, as well as oxidative and temperature stress). Under such conditions, two unusual nucleotides are being made: guanosine tetraphosphate (guanosine 5'-diphopshate 3'-diphosphate, or ppGpp) and guanosine pentaphosphate (guanosine5'-triphosphate 3'-diphosphate, or pppGpp). These nucleotides may be annotated as (p)ppGpp. Both nucleotides are synthesized by the RelA and SpoT proteins, and are being made from GDP (in case of ppGpp) or GTP (in case of pppGpp) and ATP. As of now, only the mechanism of RelA stimulation to synthesize (p)ppGpp is relatively well known. This protein becomes activated only during amino acid starvation. It binds to the ribosome and becomes active when the process of protein translation is halted due to the lack of a given charged tRNA. On the other hand, SpoT is responsible for (p)ppGpp synthesis due to starvation and stress other than amino acid starvation. Until now, it was possible to show SpoT activation only due to lipid starvation; the signals responsible for SpoT stimulation during other stresses remain unknown. SpoT possesses also another function, i.e. it removes (by hydrolysis) excess (p)ppGpp in the cell, which is very important when environmental conditions improve. It allows bacteria to rapidly switch their metabolism so as to take advantage of the newly available substrates in the most efficient way.

The (p)ppGpp nucleotide is a relatively well known transcriptional factor which directly interacts with RNA polymerase by binding to this enzyme. The place of its binding is not really clear (there are contradictory reports regarding this matter), but it probably resides near the catalytic center. Still, there is agreement that (p)ppGpp plays a major role at the stage of transcriptional initiation, i.e. at the stage when RNA polymerase binds to the promoter region on DNA, forms the closed complex, then forms the open complex (when the double stranded structure of DNA is separated and the so called transcriptional bubble is formed), and then synthesis of the first bonds in the newly synthesized RNA occurs. The stage when RNA polymerase leaves the promoter region is called transcriptional elongation, followed by termination when transcription comes to an end.

What is interesting is that (p)ppGpp exerts different effects on transcription initiating from a given promoter based on the structure of that promoter. For example, (p)ppGpp may activate or inhibit, or possibly even have just a minor effect on transcription originating from given promoter. Inhibition is generally observed in the case of ribosomal promoters (those responsible for rRNA synthesis) and tRNA promoters. Since rRNA and tRNA usually comprise

up to 80% of total cellular RNA, it was relatively easy to observe decrease in their synthesis during initial research on bacterial starvation. This tight (or stringent) correlation between amino acid starvation and decrease in rRNA and tRNA levels led to naming this type of response "the stringent response". On the other hand, in cells devoid of *relA* and *spoT* such response was not noted, and thus the processes occurring in ppGpp^0 cells (cells completely lacking (p)ppGpp) undergoing different kinds of starvations/stresses were called "the relaxed response".

Going back to the stringent response, a certain additional correlation was noted – inhibition of transcription was observed mainly for those genes whose products participate in energetically expensive processes that seem to be dispensable during starvation or stress (such as the already mentioned ribosome synthesis). On the other hand, transcription of genes whose products might be beneficial for survival is activated (for e.g. genes of amino acid biosynthesis pathways). It is known that ppGpp^0 cells are unable to grow on minimal media because of lack of such amino acid biosynthesis gene activation.

For many years, inhibitory effects of (p)ppGpp on transcription were noted under both *in vivo* and pure *in vitro* conditions. However, transcriptional activation was observed only *in vivo*; under *in vitro* conditions, activation by (p)ppGpp was observed solely with the use of cellular extracts (with one exception, where our group demonstrated activation of the paQ promoter in a pure *in vitro* system). This situation was changed dramatically in 2004, when it was discovered that the DksA protein is an indispensable cofactor for transcription activation of amino acid biosynthesis genes by (p)ppGpp. This led to an explosion of research and new challenges in the field of (p)ppGpp and stringent response.

This also became the starting point of the group of papers presented here as the 'scientific achievement'. It also became soon evident that there is a great need for a review article, which I wrote together with Dr. M. Cashel [publication #2], as this work was relatively highly cited (168 citations according to the Web of Science database).

In case of its structure, DksA is very similar to other, better known transcriptional factors, such as GreA and GreB (present in bacteria) and TFIIS (in eukaryotes). These proteins possess a very characteristic coiled coil domain, the so called "finger domain", as well as a globular domain. It is thought that these factors bind RNA polymerase through the globular domain, while the finger domain is inserted into the secondary channel of polymerase. After binding, two acidic amino acid residues at the tip of the finger domain are located near the catalytic center of polymerase and are thought to play an important role in the GreA/GreB protein function. It has been shown that both GreA and GreB affect the RNA polymerase solely at the stage of transcriptional elongation. Their role is to rescue polymerase after transcription is halted due to backtracking (in this process, RNA polymerase moves back on the DNA template such that the free 3'end of the newly synthesized RNA is out of register in the catalytic center). In order for transcription to resume, a fragment of RNA extending beyond +1 in the catalytic center must be cleaved such that the remaining RNA is realigned and a new NTP may be added. It has been shown that DksA does not possess such activity.

Because of the above, with great curiosity, I decided to test if GreA and GreB share a function with DksA, i.e. whether these proteins might play a role in the stringent response and whether they might cooperate with (p)ppGpp. My initial research approach was based on the *rrnB P1* model ribosomal promoter and was carried out both *in vivo* and *in vitro* [publication #1]. It turned out that GreA not only acts in an antagonistic manner to DksA (contrary to DksA, GreA activates transcription initiated from this promoter), but I also documented that GreA acts at the

stage of transcriptional initiation, not just elongation. The effect of GreA was independent of (p)ppGpp presence.

In case of GreB, I did not observe any substantial effect on *rrnB P1* transcription, although it does not exclude the possibility that this protein might affect other promoters.

These observations led to an idea that GreA, GreB and DksA might compete for RNA polymerase binding, which was initially confirmed by my *in vitro* observations. This would mean that there is an additional stage of transcriptional regulation, tightly correlated with the regulation exerted by (p)ppGpp. The results of these studies, carried out mainly *in vivo*, are presented in [publication #5].

It turned out that (p)ppGpp is not really necessary for transcriptional activation of amino acid biosynthesis genes to occur (earlier observations indicated that $ppGpp^+$ strains are able to grow on minimal plates, while $ppGpp^0$ strains cannot grow on such media); in order for this activation to occur it is sufficient that either DksA or GreA are overproduced, but only when the other protein is missing. Specifically, DksA overproduction facilitates growth of $ppGpp^0$ strains on minimal media only in the strains lacking *greA*; similarly, GreA overproduction facilitates growth only of $ppGpp^0$ *dksA*$^-$ strains.

Moreover, lack of growth of $ppGpp^+$ *dksA*$^-$ strains on minimal media may be rescued by GreA overproduction. This means that GreA and DksA may act interchangeably in some instances, while in others they act antagonistically. This was also confirmed by my experiments with the use of transcriptional microarrays. In addition, the data obtained in these experiments clearly demonstrated that GreB also plays a role in the network of RNA polymerase secondary channel protein interactions. Although GreB overproduction did not allow growth of $ppGpp^0$ strains on minimal medium, there is a class of genes that are regulated by this protein when the cells are devoid of *dksA*. Thanks to these observations it was possible to establish a certain hierarchy among the proteins. Namely, DksA and GreA seem to have the biggest effect and are the strongest competitors for binding to the RNA polymerase; GreB seems to be less efficient in that respect. It was also possible to establish mutual regulation of GreA, GreB and DksA that probably occurs at the level of transcription of genes coding for these proteins.

In addition, it seems that the acidic residues at the tip of the finger domain are dispensable in the regulation mentioned above, although they were earlier shown to be necessary for GreA and GreB to exert their action at the level of transcription elongation. Thus it seems plausible that these proteins are activating transcription from the amino acid biosynthesis genes by a mechanism much different from the one reported earlier.

In my work on GreA, GreB and DksA factors I also focused on regulation of genes coding for these proteins. Till then, there were no reports regarding that matter. The results of this research are also published [publication #3] – there, I showed that GreA is autoregulated. Deletion of *greA* leads to increase in *greA* promoter region activity, while GreA overproduction leads to transcription inhibition. Mapping of the *greA* promoter region indicated that there are two very strong overlapping promoters (σ^{70} dependent P1 promoter, and σ^E dependent P2 promoter). What is interesting, GreA autoinhibition that was observed *in vivo*, was not observed *in vitro*, which indicates that either the observed effect is indirect, or that an additional factor that is absent in the pure *in vitro* system is required.

Still, these experiments led to a very interesting observation – I discovered an unusual terminator in the *greA* promoter region. Not only does this terminator cause termination of about 2/3 of overall transcripts initiating from P1 and P2 (only 1/3 of transcripts reads through the

terminator and yields *greA* mRNA); but also, termination that takes place is highly imprecise: each promoter gives rise to about 10 different prematurely terminated RNA whose length differs by 1-10 nucleotides. What is more, the terminator structure is unusual as well: the hairpin structure is formed by 11 base pairs (the most common terminators possess structures made of 6-7 base pairs only).

This prompted me to take a look at the transcripts that were being formed in more detail. Once I confirmed that these transcripts are being released from the transcriptional complexes (which meant RNA polymerase is not pausing or is not stuck on DNA), I decided to test if these short RNA fragments may play any physiological role in the cell. I performed a set of experiments that clearly indicated that the arising sRNA ("short RNA") indeed affect expression of about 100 genes. Thus, the discovery of a novel sRNA - GraL ("*greA* leader") was the culmination of this work .

In the course of my research I also decided to investigate the question whether (p)ppGpp is responsible for cellular growth rate control. This problem is not a trifle since contradictory reports were being published for the past 20 years. The major points of contest were data obtained with the use of $ppGpp^0$ strains and media containing varying amounts of amino acids or carbon sources. In theory, if (p)ppGpp is responsible for growth rate control, then $ppGpp^0$ cells should contain much higher levels of cellular macromolecules (RNA and protein) than wild type cells, even when growing with similar generation times as wild type (in those cells there would be no restriction in rRNA and tRNA synthesis like in the wild type cells). In this case all experiments are performed in media containing enough nutrients so that growing cells do not run out of them and there is no cell starvation. The only limiting factor is the identity of substrates, but not the amount of substrates themselves.

Because previous reports were contradictory, it was first necessary here to establish growth conditions of $ppGpp^0$ strains in greater detail. It is known that prolonged growth of such strains leads to occurrence of spontaneous mutations in the genes coding for RNA polymerase subunits. These mutants (also called 'M+ mutants') act as if (p)ppGpp was bound with RNA polymerase at all times, even though (p)ppGpp is not present in these cells. After growth conditions minimizing M+ mutant emergence were established, and RNA and DNA levels were quantitated with the use of fluorescent dyes, it turned out that macromolecular content in $ppGpp^0$ strains indeed accumulates to much higher levels than in the wild type cells; what's more, this rate is independent of cell generation time. Polysome analysis which I performed with the use of $ppGpp^0$ and wild type cells, clearly indicated that in the case of RNA, it is the amount of ribosomes that mainly increases. Moreover, those ribosomes seem to be fully matured and functional.

Analysis of the effects of DksA indicated that this protein exerts some effect on growth rate control but only when efficiently overproduced. Thus, (p)ppGpp is clearly the major and sufficient determinant of growth rate control.

In summary, major discoveries presented in the papers comprising my scientific achievement, are:

- transcription factor GreA affects transcriptional initiation and not just elongation
- in many instances GreA and DksA act antagonistically to each other (for e.g. in the regulation of *rrnB P*1 promoter transcription), but there are also instances where they act interchangeably

- *greA* is autoregulated, and transcription of this gene occurs from two overlapping promoters (σ^{70} dependent P1 promoter, and σ^{E} dependent P2 promoter)

- *greA* promoter region contains an unusual terminator that produces novel sRNA – GraL

- mutual regulation between GreA, GreB and DksA exists, and occurs at the level of transcription of the genes coding for these proteins

- GreA, GreB and DksA compete for RNA polymerase binding and together with (p)ppGpp form a complex network responsible for global cellular regulation

-(p)ppGpp is necessary and sufficient for growth rate control to occur.

The results described above contribute to better understanding of molecular mechanisms of gene expression regulation in bacterial cells exposed to environmental stress conditions.

5. Description of other achievements

I began my Master's studies at the Intercollegiate Faculty of Biotechnology, University of Gdańsk- Medical University of Gdańsk, in 1994. I began doing research after my third year of studies, under direct supervision of Dr. Sylwia Barańska and Prof. dr. hab. Grzegorz Węgrzyn. At that time, I became interested in the field of (p)ppGpp and the effect of the stringent and relaxed response on the *oriJ* plasmid replication (publication IIA-1, Appendix 3). I had shown that similarly to λ plasmids (i.e. plasmids that contain bacteriophage λ origin of replication), *oriJ* plasmids replicate thanks to the inherited replication complex. In this case, one of the daughter plasmid copies inherits already assembled replication complex, while in order for the second daughter copy to replicate, the replication complex must be assembled *de novo*. However, there is a key difference between these two types of plasmids. Even though their replication regions possess similar structures, contrary to λ plasmids, *oriJ* plasmids undergo replication during the stringent response. These data became the basis of my Master's Thesis, which I successfully defended in June, 1999.

After finishing my studies, I entered the Ph.D. graduate program at the Biology and Oceanology Faculty, University of Gdańsk (1999-2003). I continued my research on (p)ppGpp and plasmid replication, which gave rise to publications IIA -2,3,4 (Appendix 3). With the use of *in vivo* formaldehyde crosslinking I was able to demonstrate that λ plasmid heritable replication complex is composed of λO, λP, DnaB and DnaK proteins (publication IIA -2). Composition of such complexes has not been established before. My research was facilitated by the use of *ΔrelA* strains under conditions of amino acid starvation. In this case, *de novo* replication complexes were not being assembled due to inhibition of protein synthesis, and thus it was possible to probe only pre-assembled (inherited) replication complexes.

I was also able to demonstrate that transcription initiating from the λp_O promoter plays a significant role in λ plasmid replication (publication IIA-3). In the case of plasmids carrying mutation in this promoter, replication occurs at a much lower level which is documented by lower plasmid copy number in the cells. Also, when cells undergo relaxed response (i.e. in the starved *ΔrelA* strains) and replication takes place only from inherited replication complexes, λp_O^- plasmids undergo much lower amplification than wild type plasmids. This suggests that transcription initiating from the λp_O promoter, similarly to that occurring from λp_R promoter, is

necessary for conformational changes in the *origin* region to occur and for the so-called transcriptional activation of the origin.

My next paper was related to the effects of Kid toxin on λ and *oriJ* plasmid replication (publication IIA-4). This toxin is encoded by plasmid R1 and together with the Kis antitoxin is responsible for stable maintenance of R1 plasmids in the cells. Earlier reports indicated that under *in vitro* conditions, helicase DnaB is the major target of Kid. My *in vivo* data with the use of λ and *oriJ* plasmids showed that Kid inhibits *de novo* assembly of replication complexes of these two types of plasmids. However, it did not affect replication carried out with the pre-assembled replication complexes (i.e. those inherited by one of the daughter plasmid copies). Assuming, that the major target of Kid is indeed DnaB, my data indicates that the inherited replication complex protects DnaB from Kid.

I also became interested in the effects of (p)ppGpp on transcription initiating from bacteriophage λ promoters. At that time I've spend three two-month fellowships at Dr. V. J. Hernandez laboratory (State University of New York, Buffalo, USA), where I obtained a great deal of experience in the field of *in vitro* transcription.

As a result of this collaboration, we published two papers (publications IIA-5,6 Appendix 3). One of them was related to bacteriophage λ pR promoter, where I showed that (p)ppGpp exerts its effects at many stages of transcription initiation, but the major effect was at the stage of bond formation between the first two nucleotides of the newly synthesized RNA (publication IIA-5). This work was also important because it clearly documented (p)ppGpp inhibition of transcription initiation of a promoter that forms a very stable open complex with RNA polymerase (half-life of half an hour to several hours, depending on reaction conditions). Earlier reports seemed to suggest that there is only one mechanism by which (p)ppGpp exerts its action, i.e. by decreasing half-life of the open complexes. Intrinsically unstable promoters would be then always inhibited by (p)ppGpp, and promoters that form stable open complexes with RNA polymerase, would be indirectly activated by (p)ppGpp (this hypothesis was formed before the role of DksA was discovered). As mentioned above, I showed in my work that (p)ppGpp affects many steps in transcription initiation, and I suggested that therefore there might be many different mechanisms by which (p)ppGpp promotes activation or inhibition of a given promoter. My later works on other promoters, as well as those of others, with or without use of DksA, seem to confirm that suggestion.

I also decided to investigate the effect of (p)ppGpp on transcription occurring from the λ paQ promoter (publication IIA-6, Appendix 3). Because other authors had showed earlier that this promoter is activated by (p)ppGpp *in vivo*, in my work I researched the effects of (p)ppGpp *in vitro*. This work was the first demonstration of (p)ppGpp promoter activation in a pure system. I showed that (p)ppGpp speeds up conformation change of the closed complex into an active open complex. This work was also important because it proved that (p)ppGpp activation of transcription is direct. Earlier hypothesis (popular among some circles) indicated that (p)ppGpp activation is an indirect effect of ribosomal transcription inhibition (inhibition of those promoters by (p)ppGpp would give rise to an increased pool of free RNA polymerase, which in turn would lead to increased transcription from other promoters).

The data described above were basis for my Ph.D. dissertation, which I performed under Prof. dr hab. Grzegorz Węgrzyn supervision and successfully defended in November, 2003 (title: "Regulation of transcription and replication of bacteriophage λ – the role of guanosine tetraphosphate (ppGpp) in control of promoters' activities".). After my defense, in December

2003 I became employed as an adjunct in the Department of Molecular Biology, University of Gdańsk. Then, in May 2004, I became a postdoctoral fellow at Dr M. Cashel laboratory, at the National Institutes of Health, Bethesda, USA.

Besides research and publications described in the first part of this author review, I also investigated another protein with similar structure to GreA, GreB and DksA, i.e. TraR. TraR is a protein encoded by plasmid F'. First, I developed a method of overproduction and purification of this protein, and then performed *in vitro* transcription with the use of *rrnB P1* system. It turned out that TraR, similarly to DksA, inhibits ribosomal promoters. However, surprisingly this protein does not require (p)ppGpp for inhibition to occur. Activation of amino acid biosynthesis genes by TraR also does not require (p)ppGpp presence (publication IIA-7, Appendix 3).

I also investigated regulation of the *dksA* gene expression. At about the same time, our group and the group of Dr. T. Romeo discovered autoregulation of DksA (by inhibition of *pdksA* by DksA). These data were included in publication IIA-8 (Appendix 3), where the major topic was inter-relation of the Csr system with the stringent response. My contribution also included mapping of the *dksA* promoter. These observations became valuable in further defining of interactions between GreA, GreB, DksA and (p)ppGpp.

National Institutes of Health, where I was a postdoctoral fellow, is a government research facility that does not teach students. Therefore, I did not have any students under my supervision while doing my postdoc there. However, in the later years of my postdoc, I became one of the most experienced members of the lab, and was able to advise not only postdocs in our group, but also those in neighboring labs. I believe that I became skilled in a substantial amount of laboratory techniques, and gained enough other experience to lead my own research group.

K. Potrykus 5/6/12