MONITORING OF GENES THAT RESPOND TO OVERPRODUCTION OF INSOLUBLE RECOMBINANT PROTEINS IN *ESCHERICHIA COLI* AND *BACILLUS SUBTILIS*

**Keywords:** overproduction, insoluble proteins, transcriptome, proteome, *Escherichia coli*, *Bacillus subtilis*

**Abstract.** The article summarizes and discusses recent data on the physiological response of the two industrial hosts *Escherichia coli* and *Bacillus subtilis* at the mRNA and protein level during overproduction of insoluble heterologous proteins. The molecular response of these two bacteria to overexpression of recombinant genes has been analyzed at the transcriptional level using the DNA macroarray technique and at the translational level by the two-dimensional polyacrylamide gel electrophoresis. The identification of critical process-related genes could help to find bottlenecks of such recombinant bioprocesses. Our data demonstrate that the cellular response of the Gram-positive bacterium *B. subtilis* to the overproduction of an insoluble heterologous protein is very similar to the heat shock-like response observed in the Gram-negative bacterium *E. coli*.

1. **INTRODUCTION**

*Escherichia coli* and members of the species *Bacillus* are the most frequently used prokaryotes for the industrial production of recombinant proteins. Their popularity in biotechnology is above all due to the fact that the cultivation of these organisms in large-scale bioprocesses at high cell densities is easy and usually inexpensive (Fernandez and Hoeffler, 1999). The existence of reliably operating expression systems is a prerequisite for the economical production of recombinant proteins. This was the reason why in the past the industrial production of heterologous proteins was mainly carried out by using the Gram-negative bacterium *E. coli* as it was the best known cell system and many expression systems for this organism were available. *B. subtilis* was usually applied for the production of homologous proteins, that are naturally secreted in the growth medium such as subtilisin or α-amylase (Palva, 1982; Simonen and Palva, 1993). The Gram-positive bacterium *B. subtilis* is considered as a GRAS organism (generally regarded as safe). This classification favours the use of *B. subtilis* for the production of pharmaceutical or food relevant proteins. In comparison to *E. coli*, *B. subtilis* is a more attractive host for the export of proteins into the extracellular medium, because of its naturally high secretory activity (Simoen and Palva, 1993). The secretion of target proteins makes it easier to purify the product from the host proteins or other compounds and in addition may supply better folding conditions compared to the reducing environment in the cytoplasm (Moks et al., 1987).
Examples of naturally secreted proteins of *B. subtilis* are the alkaline protease used as washing agent or amylases provided for the starch industry. At present, 60% of the commercially available enzymes are produced by *Bacillus* species. However, *E. coli* is still the most commonly used host for industrial production of heterologous proteins.

2. MONITORING OF THE PHYSIOLOGICAL STATUS OF MICROBIAL CELLS IN BIOPROCESSES

A comprehensive understanding of the physiology of the production strains under bioprocess conditions and thus a more successful directed strain optimization was impaired in the past by the lack of reliable informations on the physiological status of the cells at the molecular level during bioprocesses. The physiological status of production strains was mostly measured at the cellular level indirectly, by analyses of external variables outside the cells (Schuster, 2000). New developments in the analytic equipment now allow an analysis of microbial cells directly by measuring internal variables inside the cells. Improved physico-chemical separation techniques like chromatography or electrophoresis techniques allow an accurate separation of elemental and macromolecular components of cells. Especially the analyses of the cellular key metabolites adenine nucleotide phosphates (ATP, ADP and AMP), nicotinic acid-adenine di-nucleotide phosphates (NAD(P)H and NAD(P)) and also guanosine tetraphosphate (ppGpp) give very important information on the cellular physiology of the cells (see Neubauer and Winter this volume).

The cellular activities of the production cells are determined primarily by their RNA levels and finally at the protein level by their enzymes. Particularly the two-dimensional polyacrylamide gel electrophoresis technique (2D-PAGE), developed first by O'Farrell (1975), initiated a revolution in the molecular analyses of microbial cells. This approach enabled the first time an investigation of the protein pattern of a selected cellular compartment. A set of different proteins could be thus analyzed at one time. The improvement of this technique by various laboratories during the last ten years now allows the visualization of more than 1000 proteins on one gel (Washburn and Yates, 2000). The 2D-PAGE technique in combination with N-terminal protein sequencing or mass spectroscopic techniques, particularly MALDI-TOF (Master Assisted Laser Desorption / Ionization - Time Of Flight) mass spectrometry, enables an identification of all dominant protein spots of the investigated cellular compartments. The analyses of the proteome, which represents the sum of all proteins of one cell population or cellular compartment based on the genome sequencing and the protein identification techniques, allows the exploration of a set of bacterial genes, which are selectively induced by different environmental conditions.

Recently, the development of the DNA-array technology represents a further quantum jump in the molecular analysis of the physiological state of cells. This technique allows the analysis of the mRNA levels of all genes of a microbial genome at one time (Hauser et al., 1998). In combination with proteomics, it is the first time possible to get an almost comprehensive direct view on the physiological state of the cells.