Project: Whole Cell Model for the Infection by the Bacterial Pathogen *Pseudomonas aeruginosa* - Experiment and Bioinformatics

Dieter Jahn - **Technical University, Braunschweig - d.jahn@tu-bs.de**

Introduction

The opportunistic pathogen *Pseudomonas aeruginosa* causes severe infections of immunocompromized humans and is the leading cause of mortality in patients suffering from cystic fibrosis. *P. aeruginosa* is characterized by an intrinsic resistance to multiple antimicrobial agents and the ability to form biofilms in the environment or biofilm-like microcolonies during persistent infection of the cystic fibrosis lung [1,2]. Recent data indicates that *P.* aeruginosa adapts its metabolism to the environment of the human host. However, the regulatory and metabolic networks of infection and biofilm formation are only poorly understood. To gain insights to these complex phenomena the integration of most gene regulatory and metabolic processes to a whole cellular model is essential. This new type of approach requires the development of advanced high throughput analytical tools and sophisticated bioinformatics tools. First, high throughput and sophisticated metabolome analysis methods for *P. aeruginosa* under growth conditions present during persistent infection and biofilm mode of growth will be developed. Obtained data of the metabolome analysis will be linked with data of proteome and transcriptomic analysis using the established BRENDA and PRODORIC databases. All identified known and especially unknown open reading frames important for biofilm formation and infection will be inactivated and cellular consequences determined using the outlined methodology. The generation of a whole cell model for the regulatory circuits and metabolic capacities required during infection requires the development of new data formats, simulation and visualization of cellular networks and will allow the identification of new antibiotic targets. These targets will be verified using mouse infection experiments.

Project Status

During the first eight months of this explorative project experimental work focussed on establishing metabolite extraction protocols and investigation of metabolic patterns during different growth conditions. The bioinformatics work focussed on development of software tools for statistical analysis and comparison of the determined metabolic patterns. Moreover we constructed the first metabolic network maps and started construction of a database, which integrates all data obtained from metabolome, transcriptome and proteome analysis.

Experimental work

Different extraction protocols were tested for the isolation of metabolites from *P. aeruginosa*. Best results were obtained using a protocol based on repeated freeze thaw cycles of *P. aeruginosa* cells resuspended in methanol. This protocol was modified and tested extensively to ensure reproducibility. It allows isolation of 167 polar metabolites. GC-MS analysis identified 105 metabolites.

We started comparing metabolome patterns of *P. aeruginosa* growing in a defined glucose minimal medium during planktonic and biofilm growth. Metabolites were extracted at the midlogarithmic and early stationary phase of an anaerobically grown planktonic culture. Furthermore, metabolites were extracted from biofilm-grown *P. aeruginosa*. For that purpose we established a colony biofilm model and compared the metabolic pattern to planktonic grown cultures. Currently we complete the comparison of diverse planktonic and biofilm-grown cells by adding transcriptome and proteome data.

Bioinformatics

A database structure to store the metabolome data was modeled on a PostgreSQL database management system. Hereby the structure considers both the integration of PRODORIC [3] and BRENDA [4] data. Further on, a tool was programmed for the assignment of biochemical compounds to KEGG pathways. The 105 identified compounds were linked to 125 KEGG pathways. This assignment now serves as a basis to annotate the involved metabolic pathways. This annotation will be confirmed using *P. aeruginosa* knock-out mutants where appropriate. Two methods to analyze and classify the metabolome have been carried out. In the first method a correlation network [5] was constructed from 10 parallel probes of three different growth conditions. This analysis revealed a high correlation (> 0.995) of most metabolites in one growth condition. A total of 40 metabolites revealed correlation in all three growth conditions compared, underlining the distinct metabolic fingerprint of each growth condition. In the second method the same probes were analyzed using Kohonen Networks [6]. This completely new approach in the field of neural networks and self-organizing maps provided first convincing results in the clustering and visualization of metabolome data. Currently we test if Kohonen networks are also useful to analyze transcriptome and metabolome data.

Outlook

Currently we finish the first sets of metabolome, transcriptome and proteome data for several growth conditions. Future experimental work will also focus on different clinical isolates. For that purpose a different amino acid based medium will be tested. Moreover, we will use mutants to confirm predicted pathways and bioinformatic predictions.

A software client for an automatic integration of metabolome data is currently under development. On the basis of a new integrated comprehensive database comprising PRODORIC, BRENDA, KEGG and various other data sources a complex molecular network will be constructed. This information will be used for an improved flux balance analysis and modeling of a virtual *P. aeruginosa* cell.

*Lit.: **1.** Van Delden, C., and Iglewski, B.H. (1998) Cell-to-cell signaling and Pseudomonas aeruginosa infections. Emerg Infect Dis 4: 551-560. **2.** Davies, J.C. (2002) Pseudomonas aeruginosa in cystic fibrosis: pathogenesis and persistence. Paediatr Respir Rev 3: 128-134. **3.** Münch,R., Hiller,K., Barg,H., Heldt,D., Linz,S., Wingender,E. & Jahn,D. (2003) PRODORIC: prokaryotic database of gene regulation. Nucleic Acids Res. 31, 266-269. **4.** Schomburg,I., Chang,A., Ebeling,C., Gremse,M., Heldt,C., Huhn,G. & Schomburg D. (2004) BRENDA, the enzyme database: updates and major new developments. Nucleic Acids Res. 32, 431-433. **5.** Steuer,R., Kurths,J., Fiehn,O., Weckwerth,W. (2003) Observing and interpreting correlations in metabolomic networks. Bioinformatics 19,1019-1026. **6.** Hanke,J., Reich, J.G. (1996) Kohonen map as a visualization tool for the analysis of protein sequences: multiple alignments, domains and segments of secondary structures. Comput Appl Biosci. 12,447-454.*