

Ph.D thesis

Analysis of *Escherichia coli* amino acid transporters

Presented by Attila Szvetnik
Supervisor: Dr. Miklós Kálmán

Biology Ph.D School
University of Szeged

Bay Zoltán Foundation for Applied Research
Institute for Biotechnology

Szeged, 2008

Introduction

Membrane proteins take part in lot of basic vital processes and are indispensable parts of biomembranes. They promote the regulated transport of the required substances, mediate the transfer of environmental signals (signal-transduction), modify other molecules and create physical connection between the compartments (e.g. the cytoplasm and the extracellular space). According to this, their malfunction or absence can cause numerous illnesses. The bioinformatic analyses of the available total genome sequences revealed that membrane protein-coding genes are common, they represent the 20-30 % of ORFs. Compared to their huge importance, little is known about the structure and mechanism of membrane proteins.

Genomics, proteomics, system biology, bioinformatics and high throughput techniques can provide large amount of information, giving us the possibility to perform extensive and complex research projects. However, these efficient tools still leave unanswered questions. *Escherichia coli*, the most characterized living form in the world, has been investigated for more than 120 years. The full genome sequence of the K-12 strain was published in 1997. Nowadays, many other strains's sequence is available, nevertheless we do not have precise information about 24% of the ORF's function, and only 66% is verified experimentally. The situation of the eukaryotes is much worse, the proportion of the genes with justified function is 37% in

Saccharomyces cerevisiae, 5% in *Arabidopsis thaliana*, 4 % in *Drosophila melanogaster* and in mouse respectively.

Similar observations can be made in case of membrane proteins too. *E. coli* codes almost 900 proteins containing at least one transmembrane segment, which means that 21 % of the ORFs encode integral membrane proteins. The number of transporters is estimated to be 400, 180 with unidentified function.

In the course of my work I examined two different *Escherichia coli* amino acid transporter systems. The genes of the high affinity methionine transporter (MetD) were amongst the unidentified ORFs. We managed to prove their function with our examinations. The other part of the dissertation presents the structural analysis of the glutamate specific transporter GltS and details of its consequences.

Methods

Plasmid DNA preparation

Transformation of bacterial cells

Restrictional digestion and ligation of DNA

Transposon mutagenesis

Creation of unidirectional deletions

Polymerase chain reaction

Agarose gelelectrophoresis

Polyacrylamide gelelectrophoresis

ET-recombination

Western-blot

Measurement of alkaline phosphatase activity

Measurement of β -galactosidase activity

Aim of the study

1. Based on published information, we assumed that the high-affinity methionine transporter is encoded by the *abc-yaeE-yaeC* gene cluster of *E. coli*. Our goal was to find experimental evidence for this hypothesis.
2. The structure of the GltS glutamate permease is unknown. Our aim was to determine the membrane topology of this protein by molecular biological techniques.

Results

Identification of the high-affinity methionine transport system

In *Escherichia coli* at least two L-methionine uptake systems exist (Kadner and Watson, 1974). A high-affinity (*metD*) and a low-affinity (*metP*) transporter was characterized. The position of the *metD* system was punctually circumscribed with genetic mapping experiments, but similarly to the *metP*, its coding sequence is still unidentified.

We compared the genetic map and the genome sequence of *E. coli* and noticed, that there are only few membrane protein coding ORFs in the region of interest, and only one of them is a possible ABC-transporter. In the promoter sequence of the latter, we could identify possible binding sites of the MetJ regulator, so we could predict the function of the *abc-yaeE-yaeC* with high probability. The directed destruction of the operon indeed resulted in a D-methionine transport deficient strain. Each ORF was built in expression vectors, and the created constructions were used for complementation tests. Only the collective expression of the three ORFs could re-establish the transport.

The identification of the high-affinity methionine transport gave us the possibility to further characterize the other methionine permeases. The S-methyl-methionine transporter (MmuP) seemed to be possible candidate for the MetP system based on previously published data. To prove this assumption, we created a double deletant ($\Delta metD$, $\Delta mmuP$) methionine auxotroph *E. coli* strain. The result of the phenotype-test

clearly indicated that the MmuP is not identical with the low affinity L-methionine transporter.

Our examination of the L-methionine uptake in *E. coli* resulted in the identification of the high-affinity transport system. We revealed that the *metD* locus contains three genes. From these ORFs three proteins are expressed which form a typical ABC-transporter system. The promoter activity and phenotype tests proved that the *abc-yaeE-yaeC* gene cluster is part of the methionine regulon, according to this the genes were renamed as *metN*, *metI* and *metQ* respectively. The identification of the high-affinity transport system gives us the possibility to explore the genes of other L-methionine transporters.

Reentrant loops in GltS

With molecular biological techniques, namely reporter fusions, we analysed the membrane topology of GltS. Two reporter proteins (PhoA and LacZ) were fused to different length N-terminal parts of the GltS by gene manipulation methods. As the reporters have compartment-specific activities, after enzymatic measurements it is possible to identify the periplasmic or cytoplasmic localization of the fusion points. With large number of fusions, the position of all loop regions could be deduced.

Parallel to our study, another research team examined the topology of the GltS by the cysteine accessibility technique ((Dobrowolski és mtsai, 2007). Their structural model is in good correlation with our reporter fusion-based topology, only the 5-6. hydrophobic segments

were found to be in different position. This contradiction was resolved by assuming that the 5. predicted transmembrane segment has a sensitive structure, which is stabilized by the following 6. TMS. In the absence of the latter, and due to its special sequence composition, the 5. segment can behave as a real transmembrane segment.

Based on the combined results, GltS is a symmetrical protein. Ten transmembrane segments (TMS) were detected, the N- and C-terminal face the periplasmic side of the plasmamembrane. The protein is formed from two larger antiparallel domains containing five TMSs each. Our results are in agreement with the hypothesis that two hydrophobic reentrant loops exist between the 4-5. and 9-10. transmembrane segments. The first reentrant loop stretches into the membrane from the periplasmic space, the other from the cytoplasm. It is conceivable that the two loops are in close proximity to each other in the three dimensional structure. Thus, they can play an important role in the transport process, as it was observed with many transporters. Based on the presented two-dimensional structure, the theoretical computation of the three-dimensional structure is possible. It can also help in the better understanding of L-glutamate transport, and identification of the substrate binding sites.

Thesis points:

1. The DNA sequence upstream of the *abc* ORF behaves like a promoter.
2. The MetJ repressor can regulate the gene expression from the promoter.
3. The D-methionine transporter is encoded by the *abc-yaeE-yaeC* operon.
4. All three proteins are necessary for the transport process.
5. The low-affinity methionine transporter is not identical with the S-methyl-methionine transporter.
6. A GltS has ten transmembrane segments.
7. The N- and C-terminal are localized at the periplasmic side.
8. The protein is formed from two antiparallel domains, each of them having a possible reentrant loop.
9. In the absence of the following C-terminal GltS parts, the sequence between the 4-5. TMSs behaves like a real transmembrane segment.
10. The reporter fusions and the cysteine accessibility method gave basically compatible results.

Publications referred in the dissertation:

Szvetnik A, Gal J, Kalman M. (2007) Membrane topology of the GltS Na(+)/glutamate permease of Escherichia coli. *FEMS Microbiol Lett.* 275: 71-79.

Gal J, **Szvetnik A**, Schnell R, Kalman M. (2002) The metD D-methionine transporter locus of Escherichia coli is an ABC transporter gene cluster. *J. Bacteriol.* 184(17):4930-2.

Other publications:

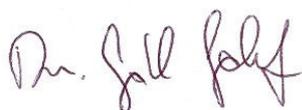
Szvetnik A. (2007) A hand-tool for picking up bacterial colonies. (Hungarian design patent)

Nagy L, Csintalan G, Kálmán E, Sipos P, **Szvetnik A.** (2003) Applications of metal ions and their complexes in medicine I *Acta Pharm Hung.* 73(4):221-36.

Társszerzői nyilatkozat

Alulírott Dr. Gál József kijelentem, hogy a Gal J, Szvetnik A, Schnell R, Kalman M. The metD D-methionine transporter locus of Escherichia coli is an ABC transporter gene cluster. J. Bacteriol. 2002. 184(17):4930-2. publikációban Szvetnik Attila meghatározó fontosságú szerepet játszott, és ezen eredményeket az „*Escherichia coli* aminosav transzporterek vizsgálata” című doktori értekezésének tézispontjaiként felhasználhatja.

Kijelentem, hogy a fenti publikációban közzétett eredményeket eddig nem használtam fel tudományos fokozat megszerzéséhez és ezt a jövőben sem teszem, illetve más jelöltnek nem adok ki hasonló jellegű nyilatkozatot a fenti publikációt illetően.



Dr. Gál József

2008.08.28. Lexington, Kentucky, Amerikai Egyesült Államok