# 学 位 論 文 の 要 旨

論文題目　　Genetic engineering of *Escherichia coli* for protein production

for functional and NMR structural study

(タンパク質の機能解析, NMR 構造解析への遺伝子改変を用いた大腸菌発現系の構築)

氏　名　　石田　洋二郎

There are a number of toxin-antitoxin (TA) systems discovered in *E. coli* and many of these TA systems have been characterized. MazF, one of the toxins from *E.coli,* functions as an ACA-specific endoribonuclease in cells. Taking advantage of MazF endoribonuclease activity in *E. coli,* a new expression system, the Single-Protein Production (SPP) system was developed.

In the SPP system, all *E. coli* cellular mRNAs are eliminated when MazF is induced. However, the mRNA of a target protein is engineered without ACA sequences while conserving the amino acid sequence so that it remains singularly intact when MazF is induced. Therefore *E. coli* is converted into a bioreactor producing only target proteins, making it especially effective for toxic proteins such as membrane proteins and antimicrobial peptides.

One of the difficult to express proteins in *E. coli* is an antimicrobial peptide (AMP) because of its toxicity to bacteria. Since the first discovery of the defensin peptide, a number of AMPs have been

isolated. One of the reasons why AMP are attractive to researchers is that while conventional antibiotics create their resistance, leading to multi-drug resistant bacteria and causing serious problems in treating inflectional diseases, AMPs do not create drug-resistant bacteria. In addition, it is relatively easy to modify the peptide sequences and possibly enhance the selectivity of its activity to bacterial cells with such changes.

Bac7(1-35) is a bovine AMP of 35 amino acid residues and is a Pro- and Arg- rich peptide. Previous studies using X-ray structure analysis have revealed one of the mechanisms of its toxicity to bacteria. Since Bac7(1-35) inhibits protein synthesis by binding to 70S ribosome, it is difficult to express in *E. coli* in its functional form.

While chemical synthesis methods have been applied to most AMP production, some recombinant expression systems such as *E. coli* and yeast systems have been successful by fusing AMP with large expression tags such as Maltose Binding Protein (MBP), Gluthathione S-Transferase (GST) and Small Ubiquitin-like Modifier protein (SUMO). However, no effective method to produce Bac7(1-35) has yet been developed.

Here I developed a novel expression system in *E.* coli, combining Protein S, a spore coat protein from *Myxococcus xanthus* with the SPP system for antimicrobial peptide production. This system produces a fusion protein which functions the same as the protein without the Protein S tag (PST). After overexpression of Bac7(1-35) using the PST-SPP system, PST-Bac7(1-35) is not only soluble, but also it functions as an antimicrobial peptide without cleaving the protein S tag from the fusion protein. This technology enables us to obtain a large amount of antimicrobial peptide in *E. coli* in a cost effective way.

I also explored expanding the capabilities of the SPP system to incorporate amino acid analogues, in particular the toxic arginine analogue, canavanine, in order to observe how protein function is altered with such substitutions. Canavanine is originally extracted from jack beans, and well known as a toxic amino acid analogue to cells. The mechanism of the toxicity is possibly due to the pKa of guanidino group in canavanine being 7 while that of arginine is 11. This charge difference causes an abnormal charge network once canavanine is incorporated into the protein leading to an eventual loss of the proper protein structures in the cells. To understand protein functions, site-directed mutagenesis has been commonly used. In addition, site-directed mutagenesis using amino acid analogues and amber codons together with modified tRNA ligase was developed to further characterize protein function. Although these systems are highly useful in studying residue specific function, a change in the protein conformation might not necessarily result in a functional change when the effect of site-directed mutagenesis is subtle. To better characterize protein function, not only site-directed residue replacement but also global replacement is necessary.

Here, the SPP system is combined with an arginine auxotrophic strain, allowing canavanine to be incorporated into protein efficiently without showing its toxicity to the cells. This is because when MazF is expressed and cell growth is arrested, the cells use canavanine for only target protein production. As an example, MazFbs, a MazF homologue from *Bacillus subtilis* which is *a* UACAU specific endoribonuclease is used as a model protein. Incorporating canavanine into MazFbs caused MazFbs(*can*) to become more helical in the structure but less stable in comparison to MazFbs because the conformation of the protein was changed. This is considered to be due to the change in the pI of MazFbs(*can*), altering the recognition sequence for cleavage to UACAUA rather than the original MazFbs UACAU recognition site. This is the first example of alteration of the RNA restriction enzyme recognition site by incorporating a toxic amino acid analogue.

Thirdly, since the auxotrophic strain is highly useful for specific amino acid labeling, I established a cost effective labeling system for NMR structural studies. Large molecular weight proteins have some dynamics, and their function and dynamics have been characterized by NMR spectrometry. However, deuteration of proteins larger than 20-kD proteins is necessary and methyl specific protonation of Ile, Leu and Val residues is commonly used to study its dynamics. For 80 kDa or larger proteins, Stereo Array Isotope Labeling (SAIL) amino acids, in which the amino acid is stereo-specifically labeled, are used as it can dramatically reduce proton density compared to proteins labeled with common labeling precursors such as α-ketoisovalerate and α-ketoisobutyrate. However, since SAIL amino acids are extremely expensive, they have not been widely used in the NMR community. Thus, the establishment of an economical method for a high expression system using the minimum amount of SAIL amino acids is highly desirable. Here, I engineered the *E.coli* strain BL21(DE3), a standard expression host for residue-, stereo- and methyl-specific labeling systems, to use minimal SAIL amino acids. As a result, I was able to reduce the usage of SAIL amino acids up to 10% compared to the standard method while maintaining protein production efficiency.

Lastly, I developed an alternative expression/labeling system for residue- stereo- methyl-specific labeled sample preparation for NMR using the common precursor, 2-acetolactate. In this system, the stereo specifically isotope-labeled 2-acetolactate is combined with genetically engineered *E. coli*, which allows proteins to be labeled in residue specific manners. Using a standard strain, Val- specific labeling is possible but Leu-specific labeling is difficult when using 2-acetolactate. To circumvent this, I engineered a biosynthetic pathway in *E. coli* to allow the cells to use 2-acetolactate for either Leu or Val synthesis so that either Leu or Val in a target protein can be labeled in a residue- and stereo-specific manner. Since auxotrophic strains were applied, the usage of the isotope labeled

2-acetolactate was reduced from 300 mg/L to 40 mg/L. To demonstrate the effectiveness of this system, I applied this method to Malate synthase G (MSG), which consists of 723 amino acid (83 kDa), Catabolite Activator Protein (CAP), a 47 kDa homodimer protein, and the peptidyl-prolyl isomerase domain of Trigger Factor, a 101 residue protein.