CHAPTER 4

DISCUSSION

To isolate protease gene from *B. subtilis* TISTR25, Tanunat (1995) had constructed a genomic library using pUC18 as a vector. Positive clones gave clear zones on skim milk-containing agar plates. One of the clone, pCSBC14, was selected for study. Its restriction map was reported elsewhere as well as in this study. Some modifications on the restriction map were reported in this study (Figure 2.1). These were the size of DNA among the restriction sites, for example, the sizes of *Eco*RI-*Hin*dIII fragments, originally reported as 0.1, 0.9 and 1.1 kb, were 0.1, 0.7 and 0.9 kb, respectively, which were confirmed by DNA sequencing data. The low protease activity of pCSBC14 clone led to the conclusion that the protease gene was mutated or not existed. To confirm the result, DNA sequencing was employed.

The 5′ and 3′ ends of pCSBC14 were conveniently sequenced by automated DNA sequencer. The DNA fragments, 0.7, 0.9 and 1.6 kb, were subcloned into M13mp18 (Figure 3.10). The 3 recombinant clones containing 0.7, 0.9 and 1.6 kb insert fragments were designated mCSBC141, mCSBC142 and mCSBC143, respectively. These clones were sequenced by manual method in both directions except mCSBC143 that was sequenced only in 5′ direction with respect to the primer used, i.e. forward primer. Moreover, the 3′end (i.e. reverse primer used) of mCSBC142 was also sequenced by automated DNA sequencer. The sequence of each clone was compared with the GenBank deposited DNA sequences using BLAST. Each of the sequence was homologous to L-glutamine D-fructose-6-phosphate amidotransferase (*gcaA*) gene of *B. subtilis* 168 in various positions except the 5′ sequence of pCSBC14 that showed no significant similarity to any sequences in the GenBank. The sequence about 700 bases in the middle of the 1.6 kb insert fragment in mCSBC143 was not determined for 2 reasons. First, the sequence was adjacent to and downstreamed from the 5′ sequence of pCSBC14. Second, the sequence was

downstreamed 3′ of the complete *gcaA* gene. Since the 5′ sequence of pCSBC14 showed no similarity to any gene in the GenBank as mentioned above, the 700 bp undetermined sequence showed not contain any genes. Each sequence was combined in the same direction by starting at the 3′end of pCSBC14. A total of 2,308 bases was determined. The sequence revealed an open reading frame of 1,803 bp, capable of encoding a protein of 600 amino acids (Figure 3.11) with high homology to *gcaA* protein of *B. subtilis* 168 (86% identity at the nucleotide sequence level). The *B. subtilis* TISTR25 protein was also homologous to the *R. meliloti*, *E. coli*, *C. albicans*, *S. cerevisiae*, human and mouse (Figure 3.14). These results of sequence comparisons suggested strongly that pCSBC14 was a clone of the structural gene for the *B. subtilis* TISTR25 L-glutamine D-fructose-6-phosphate amidotransferase (*gcaA*).

To test the activity of *B. subtilis* TISTR25 *gcaA* protein in *E. coli* DH5α harboring pCSBC14, cell extract was prepared and assayed by using the method of Ghosh *et al.* (1960). The specific activity of *gcaA* protein from *E. coli* DH5α harboring pUC18 was very low (Table 3.4). This result was supported by the work of Badet *et al.* (1987) who purified the *E. coli gcaA* protein. The specific activity of the protein form *E. coli* harboring pCSBC14 was about 36-fold higher. Thus, this result strongly indicated that the activity of *gcaA* protein from *E. coli* DH5α harboring pCSBC14 was resulted from the overexpression of this gene from pCSBC14. Since the direction of the cloned *gcaA* gene in pCSBC14 was the same as *lac* promoter, it was suspected that the expression might be driven by *lac* promoter. To test this possibility, IPTG was included in the culture medium. The result from Table 3.4 showed that the enzyme activities of cell extracts from *E. coli* DH5α harboring pCSBC14 with and without IPTG were not significant different (36.01 and 36.23 units/mg protein). Therefore, the expression of *gcaA* gene was not under the influence of *lac* promoter. It also indicated that the promoter of *gcaA* gene itself was present in the clone. From the sequence of *B. subtilis* TISTR25 *gcaA* gene, the putative promoter regions were found as TTGACT and TTGAAA at -10 and -35, respectively, which were upstreamed from the putative transcription start site (G, +1) (Figure 3.11). Besides the finding of the putative promoter, the two possible ribosome binding sites

(Shine-Dalgarno sequences) were found as AGGAGG and AGGAAG. Furthermore, the potential ρ-independent transcription terminator containing an inverted repeat, ACCCCTTT and AAAGGGGT, followed by a tract of T were also found about 114 bp from the stop codon. Recent published genome sequence of *B. subtilis* revealed that the *gcaA* gene (*glmS*) was about 200 kb from the origin of replication (*oriC*) (Kunst *et al.*, 1997).

Comparison of prokaryotic and eukaryotic *gcaA* protein sequences (Figure 3.14) revealed a relatively large region (residue 211 to 289 according to *S. cerevisiae* numbering) that was present only in the eukaryotic proteins. Since eukaryotic *gcaA* proteins differed from the bacterial enzymes in that the formers were subjected to allosteric inhibition by uridine 5′-diphosphate *N*-acetylglucosamine (UDP-GlcNAc) (Kornfeld, 1967), it seemed possible that this region contained amino acids residues that involved in the interaction with the allosteric effector.

In *E. coli*, methionine1 (at N-terminal) of *gcaA* protein was posttranslationally removed, leaving cysteine as the N-terminal residue (Badet *et al.*, 1987). This cysteine residue was highly conserved among *gcaA* proteins. It had been proposed that the N-terminal cysteine residue functioned in the glutamine amide transfer. It was very likely, from amino acid sequence homology, that the *gcaA* protein from *B. subtilis* TISTR25 would be processed similarly. The N-terminal cysteine should also function in the glutamine amide transfer.

Lysine 604 in the *E. coli gcaA* protein was proposed to be involved in the binding of D-fructose-6-phosphate (Golinelli-Pimpaneau and Badet, 1991). The Lysine 604 was in the highly conserved C-terminal region. The corresponded residue in *B. subtilis* TISTR25 *gcaA* protein was Lysine 596.

The genetic distances among nucleotide sequences and amino acid sequences of 8 organisms were calculated using Kimura 2-parameter and Dayhoff PAM matrix in Phylip version 3.5c, respectively, and used to construct phylogenetic tree. From the phylogenetic tree based on nucleotide sequence divergences (Figure 3.15), the sequence divergences between groups, i.e., the group of *B. subtilis*, yeast (*S.*

cerevisiae and *C. albicans*) and higher eukaryote (human and mouse), were higher than those in the phylogenetic tree based on amino acid sequence divergences (Figure 3.16). This was due to the degeneracy of amino acid coding sequence. Mutation in DNA may or may not result in amino acid changes. Therefore, mutations in DNA sequence were accounted for the calculated of phylogenetic tree, but the same mutations may not result in changes amino acid.

The three-dimensional structures of both *B. subtilis* TISTR25 and *E. coli gcaA* protein in Figure 3.16 were shown as N-terminal domain of about 237 amino acid residues. This domain retains the ability to bind glutamine, one of the two substrates in the synthesis of glucosamine-6-phosphate.

It was very unfortunate that there was no protease gene in pCSBC14. The pCSBC14 was probably selected as false-positive clone for protease gene. Therefore, protease genes in the genome of *B. subtilis* TISTR25 were traced again by hybridization with protease probes. Eight restriction enzymes (*Bam*HI, *Eco*RI, *Hin*dIII, *Kpn*I, *Pst*I, *Sal*I, *Sma*I and *Xba*I) were used to digest genomic DNA of *B. subtilis* TISTR25. The digests were analyzed by agarose gel electrophoresis, Southern blotting and hybridization with neutral and alkaline protease probes. Both of protease probes gave hybridization signal, but only neutral protease probe generated specific bands. Because the alkaline protease probe was relatively short (20 bases), it might give non-specific hybridization bands. In fact, more than one band were detected with this probe (Figure 3.17). Nevertheless, the results of Southern blot hybridization confirmed that both neutral and alkaline protease genes existed in the genome of *B. subtilis* TISTR25. The result also provides a basic informations on cloning the protease gene of *B. subtilis* TISTR25 which are essential for further studies.