Cloning and Sequencing to Describe the L-Glutamine D-Fructose-6-Phosphate Amidotransferase from *Bacillus subtilis* TISTR 25 Evolution Relationship and Predicted N-terminal Domain Enzyme Structure

Jarunee Vanichtanankul, Dolnapa Tanunat,

Vichien Rimphanitchayakit and Napa Siwarungson*

L-glutamine D-fructose-6-phosphate amidotransferase or glucosamine-6-phosphate synthase catalyzes the formation of glucosamine-6-phosphate and is the first rate limiting enzyme of the hexosamine biosynthetic pathway. Molecular cloning of the L-glutamine D-fructose-6-phosphate amidotransferase(*gcaA*) gene was done by partial digestion of *Bacillus subtilis* TISTR 25 chromosome with *Sau*3AI. The recombinant plasmid, pCSBC14, contains a 4.0 Kb chromosomal DNA fragment of *B. subtilis* TISTR 25 in pUC18. The pCSBC 14 was digested into 0.7, 0.9 and 1.6 Kb fragments with *Hind* III and *EcoR* I. Each of the fragments was subcloned to the M13mp18 DNA vector and transformed into *E.coli* JM109. The three subclones were sequenced by the dideoxynucleotide chain-termination method. The sequence of 2,308 bp in the 3' side of the insert in pCSBC 14 was determined. The sequence revealed an open reading frame of 1,803 bp capable of encoding a protein of 600 amino acids. This gene was compared with the GenBank deposited DNA sequences using BLAST. It showed 86% identity to *gcaA* gene of *B. subtilis* 168. The activity of the GcA enzyme in *Escherichia coli* DH5α haboring pCSBC 14 was assayed and confirmed. Therefore, the pCSBC 14 was a clone of *B. subtilis* TISTR 25 L-glutamine D-fructose-6-phosphate amidotransferase (*gcaA*) gene. From the *gcaA* sequence, two possible ribosomal binding sequence (Shine-Dalgarno sequence) were AGGAGG. The putative transcription start site was G (+1). The potential ρ-independent transcription terminator containing an inverted repeat, ACCCCTTT and AAAGGGGT, followed by a tract of T was shown. The amino acid composition was determined from the deduced amino acid sequences and compared with *Bacillus subtilis* 168 GcaA enzyme. The molecular weight of *Bacillus subtilis* TISTR 25 GcaA enzyme was calculated to be 65,431 daltons. Furthermore, the predicted amino acid sequences and nucleotide sequences of GcaA *enzyme* of *Bacillus subtilis* TISTR 25 were aligned with the GcaA enzyme of *Bacillus subtilis* 168, *Rhizobium meliloti*,

Escherichia coli, *Candida albicans*, *Saccharomyces cerevisiae*, mouse and human. The genetic distances among the nucleotide sequences of 8 organisms were calculated using Kimura 2-parameter in Phylip 3.5c and the genetic distance among amino acid sequences of 8 organisms was calculated using Dayhoff PAM matrix in Phylip 3.5c. The constructed phylogenetic tree from amino acid sequences were the same pattern as the constructed phylogenetic tree from nucleotide sequences of 8 organisms. Finally, the predicted three-dimensional structure of the N-terminal domain of *Bacillus subtilis* TISTR25 GcaA enzyme was created by using Swiss-Model at http://www.expasy.ch/Swissmod/SWISS-MODEL.html, Swiss-PdbViewer and Rasmol program. The predicted N-terminal domain three-dimensional structure of *Bacillus subtilis* TISTR 25 GcaA enzyme was compared with *Escherichia coli* GcaA enzyme in chain A. Both structures were similar. They consisted of two layers of antiparallel β pleated sheets in the middle and sandwiched by α helices.

Key words: L-glutamine D-fructose-6-phosphate amidotransferase, Glucosamine-6-phosphate synthase, gcaA gene, GcaA enzyme and *Bacillus subtilis* TISTR 25

Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand.
**Correspondence to: e-mail: siwarung@yahoo.com*

การโคลน การหาลำดับเบส คุณลักษณะ ความสัมพันธ์เชิงวิวัฒนาการ และการทำนายโดเมนทางปลาย N ของโครงสร้างเอนไซม์ของ L-Glutamine-D-Fructose-6-phosphate Amidotransferase จาก *Bacillus subtilis* TISTR 25

จารุณี วานิชธนันกูล ดลนภา ตนุนาถ วิเชียร ริมพานิชยกิจ และนภา ศิวรังสรรค์ (2548)
วารสารวิจัยวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย **30(1)**

L-glutamine D-fructose-6-phosphate amidotransferase หรือ glucosamine-6-phosphate synthase เร่งการสร้าง glucosamine-6-phosphate และเป็นปฏิกิริยาแรกที่ควบคุมวิถีการสังเคราะห์ hexosamine การโคลนระดับโมเลกุลของยีน L-glutamine D-fructose-6-phosphate amidotransferase (*gcaA*) ทำโดยการย่อยบางส่วนบนโครโมโซมของ *Bacillus subtilis* TISTR25 ด้วย *Sau*3AI พลาสมิดลูกผสม pCSBC14 ประกอบด้วยชิ้น DNA ของ *Bacillus subtilis* TISTR25 ขนาด 4.0 Kb ใส่ใน pUC18 pCSBC14 ถูกย่อยเป็นชิ้นขนาด 0.7, 0.9, 1.6 Kb ด้วย *Hind* III และ *Eco*RI ชิ้น DNA แต่ละชิ้นย่อยได้ใส่ไว้ใน M13mp18 DNA vector และส่งทอดเข้าไปใน *E.coli* JM109 subclone ทั้งสามนำมาหาลำดับเบสโดยวิธี dideoxynucleotide chain-termination พบว่าลำดับเบส 2,308 bp ทางด้าน 3' ของชิ้นที่ใส่ใน pCSBC 14 มี open reading frame ที่มีขนาด 1,803 bp สามารถถอดรหัสและแปลรหัสเป็นโปรตีนที่มีกรดอะมิโน 600 ตัว ยีนนี้เมื่อเปรียบเทียบกับที่มีในธนาคารยีนลำดับเบส DNA โดยการใช้ BLAST พบว่ามีความคล้ายคลึงกัน 86% กับยีน *gcaA* ของ *Bacillus subtilis* 168 สามารถติดตามการทำงานและระบุว่าเป็นเอนไซม์ gcaAใน *Escherichia coli* DH5α ที่มี pCSBC14 ดังนั้น pCSBC 14 เป็นโคลนหนึ่งของ *Bacillus subtilis* TISTR 25 L-glutamine D-fructose-6-phosphate amidotransferase (*gcaA*) gene จากลำดับเบสของ G*caA* จะมีลำดับเบสสำหรับการจับกับไรโบโซม เรียกว่า Shine Dalgarno sequence เป็น AGGAGG ตำแหน่งของการเริ่มต้นถอดรหัสเป็น G (+1) ตำแหน่งของสัญญาณหยุดที่ไม่ต้องการ ρ มาช่วยประกอบด้วย inverted repeat ACCCCTTT และ AAAGGGGT ติดตามด้วย T ที่ยาวเป็นสายได้แสดงไว้ องค์ประกอบของกรดอะมิโนได้ตรวจหาจากลำดับกรดอะมิโนที่แปลจากรหัสพันธุกรรม และเปรียบเทียบกับเอนไซม์ *G*caA จาก *Bacillus subtilis* 168 น้ำหนักโมเลกุลของเอนไซม์ gcaA จาก *Bacillus subtilis* TISTR 25 คำนวณเป็น 65,431 D นอกจากนั้นลำดับกรดอะมิโนที่

ทำนายของยีน *GcaA* ของ *Bacillus subtilis* TISTR25 วางเทียบกับเอนไซม์ GcaA ของ *Bacillus subtilis*168, *Rhizobium meliloti, Escherichia coli, Candida albicans, Saccharomyces cerevisiae,* มนุษย์และหนู โดยการใช้ ClustalX (1.64b) ระยะห่างยีนของลำดับเบสในสิ่งมีชีวิตทั้ง 8 ชนิด ถูกคำนวณโดย Kimura 2-parameter ใน Phylip 3.5c ช่วงของระยะห่างของยีนอยู่ระหว่าง 0.0972-1.3233 ค่าเหล่านี้ได้ใช้สำหรับการสร้าง phylogenetic tree สิ่งมีชีวิตทั้ง 8 ชนิด ได้ถูกแบ่งออกเป็น 2 กลุ่มหลัก คือ prokaryote (*Bacillus subtilis* TISTR 25, *Bacillus subtilis* 168, *Escherichia coli*, และ *Rhizobium meliloti*) และ eukaryote (*Saccharomyces cerevisiae, Candida albicans,* มนุษย์และหนู)

การประมาณระยะห่างของลำดับเบสระหว่าง 2 กลุ่มเท่ากับ 0.33088 การประมาณระยะห่างของลำดับเบสระหว่างยีน *gcaA* ของ *Bacillus subtilis* TISTR25 และ *Bacillus subtilis* 168 เป็น 0.1705 ซึ่งน้อยกว่ามากระหว่างกลุ่มของ *Bacillus subtilis* และ *Escherichia coli* เท่ากับ 0.8801 นอกจากนั้น eukaryote สามารถแบ่งแป็น 2 กลุ่ม คือ eukaryote ชั้นสูง (มนุษย์และหนู) และ eukaryote ชั้นต่ำ (*Saccharomyces cerevisiae* และ *Candida albicans*) การประมาณลำดับเบสที่ห่างกันระหว่างยีน *gcaA* ของมนุษย์และหนู เป็น 0.0972 ซึ่งน้อยกว่าระหว่าง *Candida albicans* และ *Saccharomyces cerevisiae* ที่เป็น 0.3666 ระยะทางของยีนระหว่างลำดับกรดอะมิโนของสิ่งมีชีวิตทั้ง 8 ถูกคำนวณโดยใช้ Dayhoff PAM matrix ใน Phylip 3.5c Phylogenetic tree ที่สร้างขึ้นเป็นรูปแบบเหมือนกันกับ tree ที่ถูกสร้างจากลำดับเบสของสิ่งมีชีวิตทั้ง 8 ในตอนสุดท้าย โครงสร้างสามมิติของ domain ทางปลาย N ของเอนไซม์ GcaA จาก *Bacillus subtilis* TISTR 25 ถูกสร้างจำลองโดยใช้ Swiss-Model ที่ website *http://www.expasy.ch/Swissmod/SWISS-MODEL.html,* Swiss-PdbViewer และ Rasmol program โครงสร้างสามมิติถูกเปรียบเทียบกับเอนไซม์ GcaA จาก *Escherichia coli* ในสาย A โครงสร้างทั้งสองเหมือนกันมาก ซึ่งประกอบด้วยสองชั้นของ antiparallel β pleated sheet อยู่ตรงกลางและประกบด้วย α helices

คำสำคัญ L-glutamine D-fructose-6-phosphate amidotransferase, Glucosamine-6-phosphate synthase, ยีน *gcaA* , เอนไซม์ GcaA and *Bacillus subtilis* TISTR 25

INTRODUCTION

L-glutamine D-fructose-6-phosphate amidotransferase (GcaA) or glucosamine-6-phosphate synthase (EC 2.6.1.16)+ catalyzes the formation of glucosamine-6-phosphate and is the first rate-limiting enzyme of the hexosamine biosynthetic pathway (Figure 1). It belongs to the family of amidotransferases that catalyse transfer of an amide group from glutamine to a substrate to form a new C-N bond. This enzyme is inactivated by the glutamine analogue, 6-diazo-5-oxo-L-norleucine (DON) and by iodoacetamide.[1] The eukaryotic glucosamine-6-phosphate synthase is subject to feedback inhibition by uridine 5'-diphosphate *N*-acetylglucosamine (UDP-GlcNAc), the substrate for chitin synthase in fungi, some protozoon and most invertebrates.[2] UDP-GlcNAc is a noncompetitive inhibitor with respect to both glutamine and fructose-6-phosphate, suggesting that the amido-transferase has a separate "allosteric" binding site for UDP-GlcNAc.[3] In mammalian cells, this enzyme is an insulin-regulated enzyme which controls the flux of glucose into the hexosamine pathway.

The product, glucosamine-6-phosphate ($GlcNH_2$-6-P), undergoes sequential transformations leading to the formation of UDP-GlcNAc, the major intermediate in the biosynthesis of all amino sugar containing macromolecules such as glycoproteins and chitin (β1-4 homopolymer of N-acetylglucosamine) both in prokaryotic and eukaryotic cells. Glucosamine is a common component of many macromolecules including peptidoglycan, lipopolysaccharide and teichoic acid in bacterial cell envelopes and glycolipids and glycoproteins in animal cell membranes. In higher eukaryote, glucosamine can help the body repair eroded and damaged cartilage. Since glucosamine stimulates the body's manufacture of collagen, the protein portion of the fibrous substance that holds joints together, it is used to repair damaged joints for Osteoarthritis.[4,5] Furthermore, glucosamine can be used as a precursor for the synthesis of modified nucleosides in medical treatment. For example, the *C*-nucleosides tiazofurin, showdomycin, pseudouridine, formycin, *etc*., show therapeutically useful antitumour properties. In addition, certain structurally modified nucleosides such as 5-fluorouracil, ribovirin and AZT (3'-azido-3'-deoxythymidine) are strong drugs with antiviral properties.[6]

The L-glutamine D-fructose-6-phosphate amidotransferase (glucosamine-6-phosphate synthase) is encoded by the *gcaA* gene for *B. subtilis* 168, the *glmS* gene for *Escherichia coli*[7] the *nodM* gene for *Rhizobium meliloti,*[8] the *GFA1* gene for *Saccharomyces cerevisiae*[9] and the *Candida albicans,*[10] and *GFAT* gene for human[11] and mouse.[12,13]

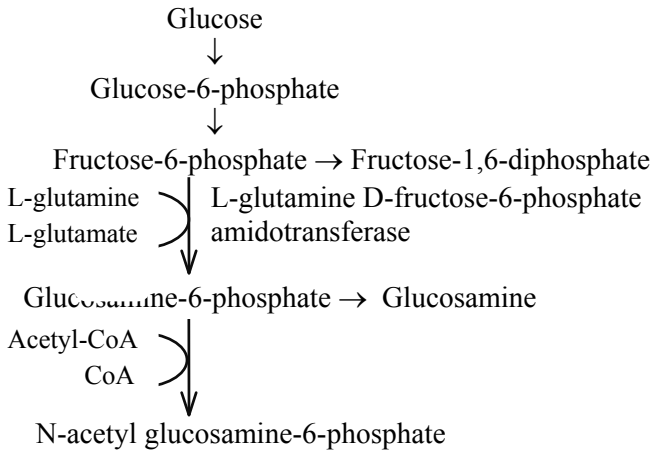


Figure 1. Hexosamine biosynthetic pathway.

MATERIALS AND METHODS

Bacterial strains, plasmids and media

Escherichia coli JM109 (F' traD36 lacI^q lacZΔM15 proA$^+$B$^+$ e14$^-$ McrA$^-$Δlac-proAB thi gryA96 Nal^r endA1 hsdR17 r^-_km^+_k relA1 supE44 recA1) and E. coli DH5α (F'/endA1 hsdR17 r^-_km^+_k supE44 thi-1 recA1 gryA Nal^r relA1 ΔlacIZYA-argF U169 deoR ϕ80 d lacΔ lacZ M15 were used as hosts. E. coli was aerobically grown in LB broth (1% Bacto-tryptone, 0.5% Bacto-yeast extract and 1% NaCl) at 37°C.

DNA isolation and manipulations

Plasmid DNA was isolated from *E.coli* by the alkaline extraction procedure.[14] *E. coli* was transformed by the electroporation method.[15] The *B. subtilis* TISTR25 genomic library consisted of DNA fragments of 2-7 kb obtained by partial *Sau*3AI digestion of *B. subtilis* TISTR25 chromosomal DNA and insertion into the *Bam*HI site of pUC18 namely pCSBC 14.[16]

DNA sequence analysis

The dideoxynucleotide chain termination method of Sanger *et al*.[17] was used with Sequenase enzyme (United State Biochemical Corp.), [α-^{35}S]dATP, 1 μg of single-stranded M13[18] or 3 μg of denatured plasmid template, and the M13-forward or reverse sequencing primers. Data handling and sequence analysis were conducted using the DNA strider 1.2 software.[19] To search for a similar sequence the BLAST[20,21] program was used. The ClastalX (1.64b) program[22] was used for sequence alignment.

Assay of L-glutamine D-fructose-6-phosphate amidotransferase

E. coli DH5α harboring pUC18 or pCSBC14 was grown at 37°C for 18 hours with vigorous aeration. The cultures were harvested by centrifuging at 4,000 rpm for 15 min, washed with 0.05 M Tris-HCl, pH 8.0 and resuspended in 0.05 M Tris-HCl, pH 8.0 (4 ml/g of packed cell). These suspensions were sonicated for 1 min on ice by using a sonicator. The extracts were centrifuged at 4,000 rpm for 15 min. The supernatant fluids were used for the enzyme assay.

Enzyme activity was measured by the method of Ghosh *et al*.[23] The assay mixture contained in a final volume of 1 ml, 20 mM fructose-6-phosphate, 15 mM L-glutamine, 37.5 mM potassium phosphate buffer, pH 7.0, 2.5 mM EDTA and cell lysate. After incubation at 37°C for 30 min, the reaction was stopped by boiling for 2 min, cooled and centrifuged. The supernatant (0.80 ml) was used for the determination of glucosamine-6-phosphate. Glucosamine-6-phosphate was determined by using a modification of the Elson-Morgan procedure described by Kenig.[24] In each experiment, two control samples, one without cell lysate and one without substrate, were assayed in the same way. The absorption reading for the control without substrate, or the sum of the readings of the two controls, was substracted from the readings obtained from the samples of the complete reaction mixtures. Solutions of glucosamine-HCl (0.1 to 1.0 mM) were assayed simultaneously for the standard curve of glucosamine. One unit of enzyme is defined as the quantity that produces 1 μmole of glucosamine-6-phosphate in 30 minutes under the condition of the assay described above.

Protein determination

Protein was determined by the method of Bradford.[25]

RESULTS

Subcloning of DNA fragment from pCSBC14

Tanunat[15] had constructed pCSBC14 which was a pUC18 recombinant plasmid, comprised of a 4.0 kb chromosomal DNA fragment of *B. subtilis* TISTR25. The three recombinant clones, comprised of 0.7, 0.9 and 1.6 kb insert fragments were named mCSBC141, mCSBC142 and mCSBC143, respectively. Each of the three clones was constructed by subcloning the fragment into M13mp18.

The RF forms of the three recombinant M13 clones were extracted by alkaline extraction and then digested with restriction endonucleases in order to determine the corrected insert fragments. The mCSBC141 and mCSBC142 were digested with both *Hin*dIII and *Eco*RI, and the mCSBC143 was digested with only *Hin*dIII. The insert fragment

sizes of mCSBC141, mCSBC142 and mCSBC143 were determined as 0.7, 0.9 and 1.6 kb, respectively (Figure 2).

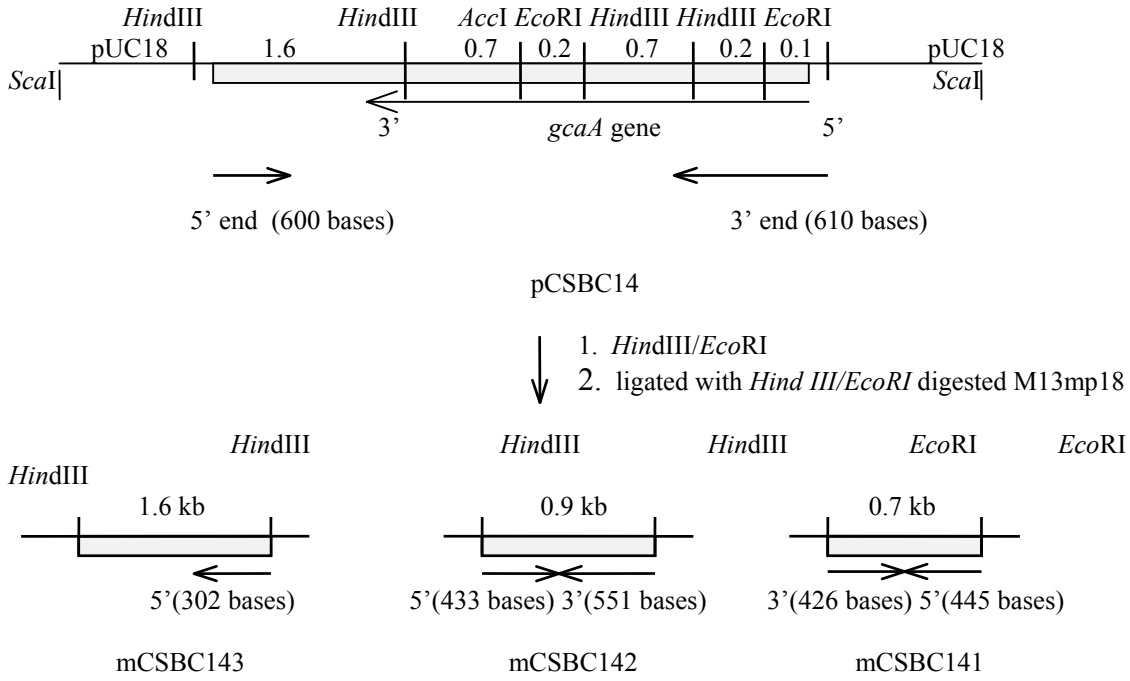


Figure 2. Sequencing orientation of individual clones. The horizontal arrows indicate the direction of forward primer and reverse primer.

DNA sequence analysis

The nucleotide sequences of the insert fragments mCSBC141, mCSBC142 and mCSBC143 were similar to *B. subtilis* 168 *gcaA* gene in various positions. The *B. subtilis* TISTR25 gene started at the 3'end of pCSBC14. Therefore, the 3' sequence of mCSBC141 and 5'sequence of mCSBC142 were inverted to a complementary strand. A DNA sequence of 2,308 bp of 4.0 kb of insert fragment of pCSBC14 was determined.

Analysis of the nucleotide sequences of pSCBC14 revealed an open reading frame of 1,803 bp. capable of encoding a protein of 600 amino acids (Figure 3). Two possible ribosomal binding sequences (Shine-Dalgarno sequence) were AGG and AGG. The putative promoter regions (-10 and -35) were TTGACT and TTGAAA, respectively. The putative transcription start site was G (+1). The potential ρ-independent transcription terminator containing an inverted repeat, ACCCCTTT and AAAGGGGT, followed by a tract of T was shown. The restriction map was created by using the DNA strider 1.2 program. The sequences of the *B. subtilis* TISTR25 gene, the upstream sequence (95 bp) and the downstream sequence (410 bp) were subjected to BLAST comparison. The results showed an 85%, 92% and 86% identity to the *B.subtilis* 168 *gcaA* gene, respectively (Figure 4). It was certain that the *B. subtilis* TISTR25 insert gene on pCSBC 14 was L-glutamine D-fructose-6-phosphate amidotransferase (*gcaA*) gene. The BLAST result of 1,803 bp of *B. subtilis* TISTR25 *gcaA* gene was obtained. The sequences of *B. subtilis* TISTR25 and *B. subtilis* 168 *gcaA* gene were shown in Figure 5. The molecular weight of *B. subtilis* TISTR25 GcaA enzyme was calculated to be 65,431 daltons. Furthermore, the predicted amino acid sequence of GcaA enzyme from *B. subtilis* TISTR25 was aligned with the GcaA enzyme of *B. subtilis* 168, *Rhizobium meliloti*,[8] *E.coli*,[7] *Candida albicans*,[10] *Saccharomyces cerevisiae*,[9] human[11] and mouse[12,13] using the ClustalX (1.64b) (Figure 6).

The genetic distances among the nucleotide sequences of 8 organisms were calculated using Kimura 2-parameter in Phylip 3.5c. The range of genetic distances was between 0.0972-1.3233 (Table 2). These values were used to construct a phylogenetic tree (Figure 7). The 8 organisms can be divided into two main groups, prokaryote (*B. subtilis* TISTR 25, *B. subtilis* 168, *E. coli* and *R. meliloti*) and eukaryote (*S. cerevisiae*, *C.* albicans, human and mouse). The estimated sequence divergence between the two groups was 0.33088. The estimated sequence divergence between the *gcaA* gene of *B. subtilis* TISTR 25 and *B. subtilis* 168 was 0.1705 which was a lot less than that between the group of *B. subtilis* and *E. coli* (0.8801). Furthermore, Eukaryote can be divided into two groups, higher eukaryote (human and mouse) and lower eukaryote (*S. cerevisiae* and *C. albicans*). The estimated sequence divergence between the two groups was 0.4344. The estimated sequence divergence between the *gcaA* gene of human and mouse was 0.0972 which was less than that between *C.albicans* and *S. cerevisiae* (0.3666).

The genetic distance among amino acid sequences of the 8 organisms was calculated using the Dayhoff PAM matrix in Phylip 3.5c. The range of genetic distances was between 0.01002-1.38757 (Table 3). These values were used to construct phylogenetic tree (Figure 8). Two groups, prokaryote and eukaryote, were also clearly seperated. The estimated sequence divergence between the two groups was 0.4355. The estimated sequence divergence between the GcaA enzyme of *B. subtilis* TISTR 25 and *B. subtilis* 168 was 0.0694.

Furthermore, the estimated sequence divergence between the higher eukaryote and lower eukaryote was 0.4618. The estimated sequence divergence between the GcaA enzyme of human and mouse was 0.01002 which was less than that between *C. albicans* and *S. cerevisiae* (0.33108). The 3D structure of *B.subtilis* TISTR 25 GcaA enzyme was compared with *E.coli* GcaA enzyme in chain A (Figure 9). Both structures were similar. They consisted of two layers of antiparallel β pleated sheets in the middle sandwiched by α helices.

```
                                    -35              -10       +1                ******
  1  GATCTTCCAAAAAAACATGTGGGAGGGGACGATTGAAAGTCCCCTTGTAATTTGACTTTCTTCGTCTCCTTTTGCACCTTTAGGAGGAAG

           M  C  G  I  V  G  Y  I  G  Q  L  D  A  K  E  I  L  L  K  G  L  E  K  L  E  Y  R  G
 91  AAAATATGTGTGGAATTGTAGGTTACATCGGTCAGCTTGATGCGAAAGAGATTTTGTTAAAAGGATTAGAGAAGCTTGAGTACCGCGGTT
                                                                          HindIII
     Y  D  S  A  G  I  A  V  A  N  E  Q  G  V  H  V  Y  K  E  K  G  R  I  A  D  L  R  E  V  V
181  AAAACTCTGCCGGTATCGCTGTGGCGAATGAGCAGGGCGTGCATGTGTACAAAGAAAAAGGCCGCATCGCCGACCTTCGTGAAGTGGTGG

     D  H  T  V  E  S  Q  A  G  I  G  H  T  R  W  A  T  H  G  E  P  S  F  L  N  A  H  P  H  Q
271  ATCACACGGTTGAATCTCAAGCGGGAATCGGCCATACACGCTGGGCGACTCACGGTGAACCAAGCTTCCTGAACGCTCACCCGCATCAAA
                                                          HindIII
     S  A  L  G  R  F  T  L  V  H  N  G  V  I  E  N  Y  V  Q  L  K  R  E  Y  L  E  N  V  E  L
361  GCGCACTCGGCCGCTTTACACTTGTTCACAATGGTGTGATCGAGAACTATGTTCAGCTGAAGCGCGAATATCTTGAAAACGTTGAACTGA

     K  S  D  T  D  T  E  V  V  V  Q  M  I  E  Q  F  V  A  G  G  L  S  T  E  E  A  F  R  K  T
451  AAAGCGACACGGACACTGAAGTAGTCGTTCAAATGATCGAGCAATTTGTGGCGGGAGGACTCAGCACAGAAGAAGCGTTCCGCAAAACAC

     L  T  L  L  K  G  S  Y  A  I  A  L  F  D  G  E  N  T  D  T  I  Y  V  A  K  N  K  S  P  L
541  TGACTCTGTTAAAAGGCTCTTACGCAATTGCATTATTTGACGGTGAAAACACAGACACCATTTACGTTGCAAAAAACAAAAGCCCTCTGT

     L  I  G  L  G  D  T  F  N  V  V  A  S  D  A  M  A  M  L  Q  V  T  N  E  Y  V  E  L  L  D
631  TAATCGGCCTTGGAGATACGTTTAACGTCGTGGCATCTGACGCGATGGCTATGCTTCAAGTAACGAATGAATACGTTGAGCTTTTGGACA

     K  E  M  V  I  V  T  K  D  E  A  V  I  K  N  L  D  G  E  V  M  T  R  A  S  Y  I  A  E  L
721  AAGAAATGGTGATCGTGACAAAAGATGAAGCCGTGATTAAAAACCTTGACGGTGAAGTCATGACACGTGCGTCTTATATCGCTGAGCTTG

     D  A  S  D  I  E  K  G  T  Y  P  H  Y  M  L  K  E  T  D  E  Q  P  L  V  M  R  K  I  I  Q
811  ACGCCAGTGATATCGAAAAAGGCACATACCCTCACTACATGTTAAAAGAAACGGATGAGCAGCCGCTTGTTATGCGCAAAATCATCCAAA

     T  Y  Q  D  E  N  G  R  L  A  V  A  G  D  V  A  D  A  V  A  E  A  D  R  I  Y  I  V  A  C
901  CGTATCAGGACGAAAACGGCAGACTGGCCGTGGCCGGCGATGTCGCTGACGCCGTGGCGGAAGCGGACCGCATTTATATCGTGGCTTGCG

     G  T  S  Y  H  A  G  L  V  G  K  Q  Y  I  E  M  W  A  N  V  P  V  E  V  H  V  A  S  E  F
991  GAACGAGCTACCACGCCGGTCTTGTCGGGAAACAATATATTGAAATGTGGGCAAACGTACCGGTTGAAGTGCATGTAGCGAGTGAATTCT
                                                                                    EcoRI
     S  Y  N  M  P  L  L  S  K  K  P  L  F  I  F  L  S  Q  S  G  E  T  A  D  S  R  A  V  L  V
1081 CTTACAACATGCCGCTTCTGTCTAAGAAGCCGCTCTTTATCTTCCTTTCTCAAAGCGGAGAAACAGCGGACAGCCGCGCCGTGCTTGTTC

     Q  V  K  A  L  G  H  K  A  L  T  I  T  N  V  P  G  S  T  L  S  R  E  A  D  Y  T  L  L  L
1171 AAGTCAAAGCGCTGGGTCACAAAGCGCTGACGATTACAAACGTTCCGGGATCAACGCTTTCCCGTGAAGCGGATTACACATTGCTTCTGC

     H  A  G  P  E  I  A  V  A  S  T  K  A  Y  T  A  Q  I  A  V  L  A  I  L  A  S  V  A  A  E
1261 ACGCAGGCCCTGAGATCGCCGTGGCATCAACAAAAGCGTATACGGCTCAGATTGCCGTCCTCGCGATCCTTGCGTCCGTTGCAGCAGAAC
                                          AccI
     R  N  G  V  D  I  G  F  D  L  V  K  E  L  G  I  A  A  N  A  M  E  A  L  C  D  Q  K  D  E
1351 GCAACGGCGTTGATATCGGTTTTGATTTAGTCAAAGAATTAGGTATCGCGGCAAACGCCATGGAAGCCCTCTGCGACCAGAAGGACGAAA
```

```
     M  E  M  I  A  R  E  Y  L  T  V  S  R  N  A  F  F  I  G  R  G  L  D  Y  F  V  C  V  E  G
1441 TGGAAATGATCGCACGTGAGTACCTGACTGTTTCAAGAAACGCTTTCTTCATCGGCCGCGGCCTTGACTACTTCGTGTGTGTCGAAGGCG

     A  L  K  L  K  E  I  S  Y  I  Q  A  E  G  F  A  G  G  E  L  K  H  G  T  I  A  L  I  E  E
1531 CCCTGAAGCTGAAAGAGATTTCTTACATCCAGGCGGAAGGCTTCGCCGGCGGCGAGCTGAAGCATGGAACAATCGCTCTGATTGAAGAAG

     G  T  P  V  F  A  L  A  T  Q  E  H  V  N  L  S  I  R  G  N  V  K  E  V  A  A  R  G  A  N
1621 GAACACCGGTCTTTGCGCTTGCGACACAAGAACACGTCAACCTGAGCATCCGCGGTAATGTGAAGGAAGTCGCAGCCCGCGGCGCCAACA

     T  C  I  I  S  L  K  G  L  E  D  A  D  D  R  F  I  L  P  E  V  N  P  A  L  R  P  L  V  S
1711 CTTGCATCATCTCGCTGAAAGGCTTAGAAGACGCAGACGACAGATTCATCCTGCCGGAAGTCAACCCTGCGCTTCGTCCGCTGGTTTCTG

     V  V  P  L  Q  L  I  A  Y  Y  A  A  L  H  R  G  C  D  V  D  K  P  R  N  L  A  K  S  V  T
1801 TTGTGCCATTGCAGCTGATCGCTTACTACGCTGCACTGCACCGCGGCTGTGACGTTGATAAACCGCGCAACCTTGCGAAGAGTGTTACGG

     V  E ---
1891 TGGAATAATATATTTAACCCCTTTGGTTATTAGTGTGAATTTAAAATAAAGTCGCGATGTTTATAAAAAAGATGCGATGTTTAAAATAAA

1981 GTCGCGANGTTTATAAAAAAGTTGCGAAGCTTACCCCTTTGGATATGATTATCTAAAGGGGTGTTTTTGTGTCGAAAAGAAAAAGAACA
                           HindIII
2071 TCTAAAATTGATAAGTGGATTAAAGAGGGTCGAGGAACTGGCAGTGGGGCATGATTATCAAGATCTTTCCTCATTAGGTCGGTCAACAAG

2161 ATTAAAAGGTATAAAAACCGGCAGACAACATGAGTTTTTATCGGATTTGGAACGAAACTACTTTTATTTAACTGAATTTTCTGATGT

2251 TATTTTAGATATTCGTGAACAATTTCCTTTATTACCACAAGAAGAGACGTTTGCCATTGCTG
```

Figure 3. Nucleotide sequence and deduced amino acid sequence of *B. subtilis* TISTR25 *gcaA* gene. Two possible Shine-Dalgarno sequences (****) and a possible promoter (-35 and -10 regions) were indicated. The putative ρ-independent termination codon (---) was indicated.

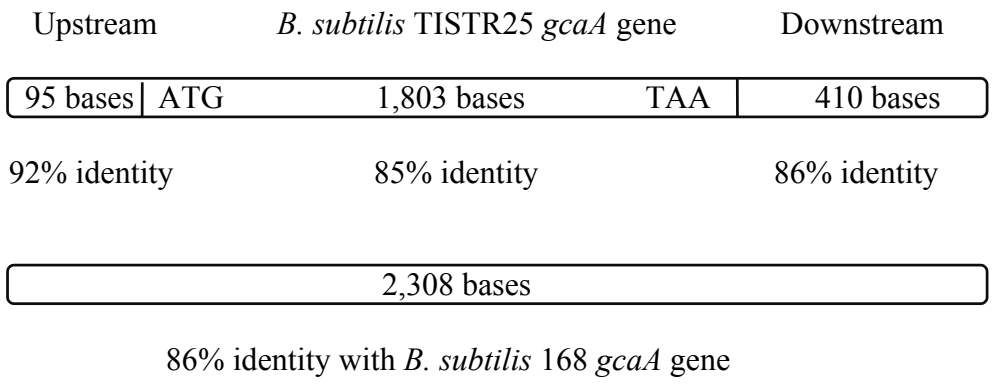


Figure 4. Percent identity of *B. subtilis* TISTR25 and *B. subtilis* 168 *gcaA* gene by BLAST comparison.

```
B.subtilis TISTR25      ------------------------------------------------------------
B.subtilis 168          AAACAAGCTGTCTGTTCTTATTTCTCAATAGGAAAGAAGACGGATTATTCCTTACCTATA

B.subtilis TISTR25      ------------------------------------------------------------
B.subtilis 168          ATTATAGCGCCCGAACTAAGCGCCCGGAAAAAGGCTTAGTTGACGAGGATGGAGGTTATC

B.subtilis TISTR25      ------------------------------------------------------------
B.subtilis 168          GAATTTCGCGGATCCTCCCGGCTGAGTGTGCAGATCACAGCCGTAAGAATTTCTTCAAAC

B.subtilis TISTR25      ----------------------------------GATCTTCCAAAAAAACATGTGGGA
B.subtilis 168          CAAGGGGGTGACTCCTTGAACAAAGAGAAATCACATGATCTTCCAAAAAA-CATGTAGGA
                                                           ************** ***** ***

B.subtilis TISTR25      GGGGACGATTGAAAGTCCCCTTGTAATTTGACTTTCTTCGTCTCCTTTTGCACCTTTAGG
B.subtilis 168          GGGGACGATTGAAAGTCCCCTTGAAATTTGACTTTCTTCGTCTCCTTTTACAATCTTAGG
                        ********************** ************************ **   *****

B.subtilis TISTR25      AGGAAGAAAA-TATGTGTGGAATTGTAGGTTACATCGGTCAGCTTGATGCGAAAGAGATT
B.subtilis 168          AGGAAGAAAAATATGTGTGGAATCGTAGGTTATATCGGTCAGCTTGATGCGAAGGAAATT
                        ********** ************ ******** ******************** ** ***

B.subtilis TISTR25      TTGTTAAAAGGATTAGAGAAGCTTGAGTACCGCGGTTATGACTCTGCCGGTATCGCTGTG
B.subtilis 168          TTATTAAAAGGGTTAGAGAAGCTTGAGTATCGCGGTTATGACTCTGCTGGTATTGCTGTT
                        ** ******** ***************** ***************** ***** *****

B.subtilis TISTR25      GCGAATGAGCAGGGCGTGCATGTGTACAAAGAAAAAGGCCGCATCGCCGACCTTCGTGAA
B.subtilis 168          GCCAACGAACAGGGAATCCATGTGTTCAAAGAAAAAGGACGCATTGCAGATCTTCGTGAA
                        ** ** ** *****  * ******* ************ ***** ** ** *********

B.subtilis TISTR25      GTGGTGGATCACACGGTTGAATCTCAAGCGGGAATCGGCCATACACGCTGGGCGACTCAC
B.subtilis 168          GTTGTGGATGCCAATGTAGAAGCGAAAGCCGGAATTGGGCATACTCGCTGGGCGACACAC
                        ** ******  **  ** *** *  **** ***** ** ***** *********** ***

B.subtilis TISTR25      GGTGAACCAAGCTTCCTGAACGCTCACCCGCATCAAAGCGCACTCGGCCGCTTTACACTT
B.subtilis 168          GGCGAACCAAGCTATCTGAACGCTCACCCGCATCAAAGCGCACTGGGCCGCTTTACACTT
                        ** **********  ***************************** ***************

B.subtilis TISTR25      GTTCACAATGGTGTGATCGAGAACTATGTTCAGCTGAAGCGCGAATATCTTGAAAACGTT
B.subtilis 168          GTTCACAACGGCGTGATCGAGAACTATGTTCAGCTGAAGCAAGAGTATTTGCAAGATGTA
                        ******** ** ***************************  ** *** *  ** * **

B.subtilis TISTR25      GAACTGAAAAGCGACACGGACACTGAAGTAGTCGTTCAAATGATCGAGCAATTTGTGGCG
B.subtilis 168          GAGCTCAAAAGTGACACCGATACAGAAGTAGTCGTTCAAGTAATCGAGCAATTCGTCAAT
                        ** ** ***** ***** ** ** *************** * *********** **

B.subtilis TISTR25      GGAGGACTCAGCACAGAAGAAGCGTTCCGCAAAACACTGACTCTGTTAAAAGGCTCTTAC
B.subtilis 168          GGAGGACTTGAGACAGAAGAAGCGTTCCGCAAAACACTTACACTGTTAAAAGGCTCTTAT
                        ********    ************************** ** *****************

B.subtilis TISTR25      GCAATTGCATTATTTGACGGTGAAAACACAGACACCATTTACGTTGCAAAAAACAAAAGC
B.subtilis 168          GCAATTGCTTTATTCGATAACGACAACAGAGAAACGATTTTTGTAGCGAAAAACAAAAGC
                        ******** ***** **    ** **** *** ** ****  ** ** ************

B.subtilis TISTR25      CCTCTGTTAATCGGCCTTGGAGATACGTTTAACGTCGTGGCATCTGACGCGATGGCTATG
B.subtilis 168          CCTCTATTAGTAGGTCTTGGAGATACATTCAACGTCGTAGCATCTGATGCGATGGCGATG
                        ***** *** * ** *********** ** ******** ******** ******** ***

B.subtilis TISTR25      CTTCAAGTAACGAATGAATACGTTGAGCTTTTGGACAAAGAAATGGTGATCGTGACAAAA
B.subtilis 168          CTTCAAGTAACCAACGAATACGTAGAGCTGATGGATAAAGAAATGGTTATCGTCACTGAT
                        *********** ** ******** *****  **** *********** ***** **  *

B.subtilis TISTR25      GATGAAGCCGTGATTAAAAACCTTGACGGTGAAGTCATGACACGTGCGTCTTATATCGCT
B.subtilis 168          GACCAAGTTGTCATCAAAAACCTTGATGGTGACGTGATTACACGTGCGTCTTATATTGCT
                        **  ***  ** ** *********** ***** ** ** ***************** ***

B.subtilis TISTR25      GAGCTTGACGCCAGTGATATCGAAAAAGGCACATACCCTCACTACATGTTAAAAGAAACG
B.subtilis 168          GAGCTTGATGCCAGTGATATCGAAAAAGGCACGTACCCTCACTACATGTTGAAAGAAACG
                        ******** ********************** ***************** *********

B.subtilis TISTR25      GATGAGCAGCCGCTTGTTATGCGCAAAATCATCCAAACGTATCAGGACGAAAACGGCAGA
B.subtilis 168          GATGAGCAGCCTGTTGTTATGCGCAAAATCATCCAAACGTATCAAGATGAAAACGGCAAG
                        ***********  ****************************** ** **********
```

```
B.subtilis TISTR25      CTGGCCGTGGCCGGCGATGTCGCTGACGCCGTGGCGGAAGCGGACCGCATTTATATCGTG
B.subtilis 168          CTGTCTGTGCCTGGCGATATCGCTGCCGCTGTAGCGGAAGCGGACCGCATCTATATCATT
                        *** * *** * ****** ****** *** ** ***************** ****** *

B.subtilis TISTR25      GCTTGCGGAACGAGCTACCACGCCGGTCTTGTCGGGAAACAATATATTGAAATGTGGGCA
B.subtilis 168          GGCTGCGGAACAAGCTACCATGCAGGACTTGTCGGTAAACAATATATTGAAATGTGGGCA
                        *  ******** ******** ** ** ******** ************************

B.subtilis TISTR25      AACGTACCGGTTGAAGTGCATGTAGCGAGTGAATTCTCTTACAACATGCCGCTTCTGTCT
B.subtilis 168          AACGTGCCGGTTGAAGTGCATGTAGCGAGTGAATTCTCCTACAACATGCCGCTTCTGTCT
                        ***** ********************************* ********************

B.subtilis TISTR25      AAGAAGCCGCTCTTTATCTTCCTTTCTCAAAGCGGAGAAACAGCGGACAGCCGCGCCGTG
B.subtilis 168          AAGAAACCGCTCTTCATTTTCCTTTCTCAAAGCGGAGAAACAGCAGACAGCCGCGCGGTA
                        ***** ******** ** ************************* *********** **

B.subtilis TISTR25      CTTGTTCAAGTCAAAGCGCTGGGTCACAAAGCGCTGACGATTACAAACGTTCCGGGATCA
B.subtilis 168          CTCGTTCAAGTCAAAGCGCTCGGACACAAAGCCCTGACAATCACAAACGTACCTGGATCA
                        ** ***************** ** ******** ***** ** ******** ** ******

B.subtilis TISTR25      ACGCTTTCCCGTGAAGCGGATTACACATTGCTTCTGCACGCAGGCCCTGAGATCGCCGTG
B.subtilis 168          ACGCTTTCTCGTGAAGCTGACTATACATTGCTGCTTCATGCAGGCCCTGAGATCGCTGTT
                        ******** ******** ** ** ******** ** ** ***************** **

B.subtilis TISTR25      GCATCAACAAAAGCGTATACGGCTCAGATTGCCGTCCTCGCGATCCTTGCGTCCGTTGCA
B.subtilis 168          GCGTCAACGAAAGCATACACTGCACAAATCGCAGTTCTGGCGGTTCTTGCTTCTGTGGCT
                        ** ***** ***** ** ** ** ** ** ** ** ** *** * ***** ** ** **

B.subtilis TISTR25      GCAGAACGCAACGGCGTTGATATCGGTTTTGATTTAGTCAAAGAATTAGGTATCGCGGCA
B.subtilis 168          GCTGACAAAAATGGCATCAATATCGGATTTGACCTCGTCAAAGAACTCGGTATCGCTGCA
                        ** **    ** *** *  ******* *****  * ********* * ******** ***

B.subtilis TISTR25      AACGCCATGGAAGCCCTCTGCGACCAGAAGGACGAAATGGAAATGATCGCACGTGAGTAC
B.subtilis 168          AACGCAATGGAAGCTCTATGCGACCAGAAAGACGAAATGGAAATGATCGCTCGTGAATAC
                        ***** ******** ** *********** ******************** ***** ***

B.subtilis TISTR25      CTGACTGTTTCAAGAAACGCTTTCTTCATCGGCCGCGGCCTTGACTACTTCGTGTGTGTC
B.subtilis 168          CTGACTGTATCCAGAAATGCTTTCTTCATCGGACGCGGCCTTGACTACTTCGTATGTGTC
                        ******** ** ***** ************** ******************** ******

B.subtilis TISTR25      GAAGGCGCCCTGAAGCTGAAAGAGATTTCTTACATCCAGGCGGAAGGCTTCGCCGGCGGC
B.subtilis 168          GAAGGCGCACTGAAGCTGAAAGAGATTTCTTACATCCAGGCAGAAGGTTTTGCCGGCGGT
                        ******** ********************************* ***** ** ********

B.subtilis TISTR25      GAGCTGAAGCATGGAACAATCGCTCTGATTGAAGAAGGAACACCGGTCTTTGCGCTTGCG
B.subtilis 168          GAGCTAAAGCACGGAACGATTGCCTTGATCGAACAAGGAACACCAGTATTCGCACTGGCA
                        ***** ***** ***** ** **  **** *** ********** ** ** ** ** **

B.subtilis TISTR25      ACACAAGAACACGTCAACCTGAGCATCCGCGGTAATGTGAAGGAAGTCGCAGCCCGCGGC
B.subtilis 168          ACTCAAGAGCATGTAAACCTAAGCATCCGCGGAAACGTCAAAGAAGTTGCTGCTCGCGGA
                        ** ***** ** ** ***** *********** ** ** ** ***** ** ** *****

B.subtilis TISTR25      GCCAACACTTGCATCATCTCGCTGAAAGGCTTAGAAGACGCAGACGACAGATTCATCCTG
B.subtilis 168          GCAAACACATGCATCATCTCACTGAAAGGCCTAGACGATGCGGATGACAGATTCGTATTG
                        ** ***** *********** ********* **** ** ** ** ********* *  **

B.subtilis TISTR25      CCGGAAGTCAACCCTGCGCTTCGTCCGCTGGTTTCTGTTGTGCCATTGCAGCTGATCGCT
B.subtilis 168          CCGGAAGTAAACCCAGCGCTTGCTCCGTTGGTATCTGTTGTTCCATTGCAGCTGATCGCT
                        ******** ***** ******  **** **** ******** ******************

B.subtilis TISTR25      TACTACGCTGCACTGCACCGCGGCTGTGACGTTGATAAACCGCGCAACCTTGCGAAGAGT
B.subtilis 168          TACTATGCTGCACTGCATCGCGGCTGTGATGTGGATAAACCTCGTAACCTTGCGAAGAGT
                        ***** *********** *********** ** ******** ** ***************

B.subtilis TISTR25      GTTACGGTGGAATAATATAT--TTAACCCCTTTGGTTA---TTA-------GTGTGAATT
B.subtilis 168          GTTACTGTGGAGTAATAAATGTTTAACCCCTTTGGATAAGATTATCTAAAGGGGTGTTTT
                        ***** ***** ***** **  ************* **   ***       * ***  **

B.subtilis TISTR25      TAA----AATAAAGTCGCGATGTTTATA----AAAAA--GATGC--GATGTTTAAAATAA
B.subtilis 168          TATGTCCAAAAGCGAAAGAACGTCTACAGTGGAAAAATGGATTAAAGAAGTTGATGGTAC
                        **     ** *  *     * ** ** *    *****  ***    ** *** *   **

B.subtilis TISTR25      AG----------TCG--CGANGTTTATAAAAAAGTTGCGAAGCTTACCCCTTTGGATATG
B.subtilis 168          AGGACCCGACTATCGATCGACCATGGCTATAGATTTAAGATG-TTTCCTCATTAGATCGG
                        **          ***  ***   *    * * * **  ** * ** ** * ** ***  *
```

```
B.subtilis TISTR25      ATTATC---TAAAGGGGTGTT-----TTTGTGTCGAAAAGAAAAAGAACATCTAAAATTG
B.subtilis 168          TCAATCAGGTTAAAAGGCATTAAAACTTGGAGAGAAACTGCTTTTGCCTAACTGAGTTTT
                           ***   * **  **  **     ** * *   **  *     *   * ** *  **

B.subtilis TISTR25      ATAAGTGGATTAAAGAGGGTCGAG-GAACTGGC--AGTGGGGCATGATTATCAAGATCTT
B.subtilis 168          CTAATGTTAATACAGATATAAGAGAGCAGTTCCTCATTAAGATACGCTTTTGATAGCAGT
                         ***    * ** ***     *** * * *  *  * * *  *  * * ** * *      *

B.subtilis TISTR25      TCCTCATTAGGTCGGTCAACAAGATTAAAAGGTATAAAA-ACCGGCA-GACAACATGAGT
B.subtilis 168          TGGAATTTAGCATGCTAGAAATCCGAAAACCGGAGAGCCTATTGGAATGATAACGTGTTT
                        *     ****   * *  * *     ***  * * *    *  ** * ** *** **  *

B.subtilis TISTR25      TTTTATCGG--ATTTGGAACGAAACTACTTTTATTTAACTGAATTTTCTGATGTTATTTT
B.subtilis 168          TTTTATTGACAATTAATAAAGGGCAAAGAATTCTAGAAGTCGCACTTACAATGTACTTGT
                        ****** *   ***   ** *     *   ** *  ** *     **   ****  ** *

B.subtilis TISTR25      AGATATTCGTGAACAATTTCCTTTATTACCACAAGAAGAGACGTTTGCCATTGCTG----
B.subtilis 168          GGGATTAGGTTTTAAAGTTCATTCTTTTTTTATTGAGTACATTCCGGTTAAAGTTGAAAT
                         *   *  **    ** *** **  **       **  * *     *  *  * **

B.subtilis TISTR25      ---
B.subtilis 168          GAT
```

Figure 5. Sequence alignment of *B. subtilis* TISTR25 and *B. subtilis* 168 *gcaA* gene by using ClustalX (1.64b). Asterisk (*) indicates identity.

Table 1. Amino acid composition of *B. subtilis* TISTR25 GcaA enzyme in comparison with *B. subtilis* 168 GcaA enzyme.

Amino acids	*B. subtilis* TISTR25 GcaA enzyme				*B. subtilis* 168 GcaA enzyme			
	n	n(%)	MW	MW(%)	n	n(%)	MW	MW(%)
A ala alanine	67	11.2	4759	7.3	66	11.0	4688	7.2
C cys cysteine	6	1.0	618	0.9	6	1.0	618	0.9
D asp aspartic acid	32	5.3	3680	5.6	35	5.8	4025	6.2
E glu glutamic acid	51	8.5	6581	10.1	46	7.7	5935	9.1
F phe phenylalanine	16	2.7	2353	3.6	17	2.8	2500	3.8
G gly glycine	45	7.5	2565	3.9	45	7.5	2565	3.9
H his histidine	15	2.5	2055	3.1	14	2.3	1918	2.9
I ile isoleucine	38	6.3	4297	6.6	40	6.7	4523	6.9
K lys lysine	32	5.3	4099	6.3	34	5.7	4355	6.7
L leu leucine	62	10.3	7011	10.7	60	10.0	6785	10.4
M met methionine	13	2.2	1703	2.6	12	2.0	1572	2.4
N asn asparagine	22	3.7	2508	3.8	25	4.2	2851	4.4
P pro proline	16	2.7	1552	2.4	17	2.8	1649	2.5
Q gln glutamine	19	3.2	2433	3.7	22	3.7	2817	4.3
R arg arginine	23	3.8	3590	5.5	20	3.3	3122	4.8
S ser serine	28	4.7	2436	3.7	27	4.5	2349	3.6
T thr threonine	32	5.3	3233	4.9	30	5.0	3031	4.6
V val valine	57	9.5	5646	8.6	59	9.8	5845	9.0
W trp tryptophan	2	0.3	372	0.6	2	0.3	372	0.6
Y tyr tyrosine	24	4.0	3913	6.0	23	3.8	3750	5.7
Total	600	100	65431	100	600	100	65296	100

```
B.subtilisTISTR25  MCGIVGYIG------QLDAKEILLKGLEKLEYRGYDSAGIAVANEQ------------GV   42
B.subtilis168      MCGIVGYIG------QLDAKEILLKGLEKLEYRGYDSAGIAVANEQ------------GI   42
R.meliloti         MCGIVGIVG------HQPVSERLVEALEPLEYRGYDSAGVATMDAG------------TL   42
E.coli             MCGIVGAIA------QRDVAEILLEGLRRLEYRGYDSAGLAVVDAEG-----------HM   43
S.cerevisiae       MCGIFGYCNYLVERSRGEIIDTLVDGLQRLEYRGYDSTGIAIDG----------DEADST   50
C.albicans         MCGIFGYVNFLVDKSRGEIIDNLIEGLQRLEYRGYDSAGIAVDGKLTKDPSNGDEEYMDS   60
Human              MCGIFAYLNYHVPRTRREILETLIKGLQRLEYRGYDSAGVGFDGGNDKD---WEANACKT   57
Mouse              MCGIFAYLNYHVPRTRREILETLIKGLQRLEYRGYDSAGVGLDGGNDKD---WEANACKI   57
                   ****..         :    : *:..*. ********:*:.  .

B.subtilisTISTR25  HVYKEKGRIADLREVVDHT--------VESQAGIGHTRWATHGEPSFLNAHPHQSALG-R   93
B.subtilis168      HVFKEKGRIADLREVVDAN--------VEAKAGIGHTRWATHGEPSYLNAHPHQSALG-R   93
R.meliloti         QRRRAEGKLGNLREKLKEAP-------LSGTIGIAHTRWATHGAPTERNAHPHFTEG---   92
E.coli             TRLRRLGKVQMLAQAAEEHP-------LHGGTGIAHTRWATHGEPSEVNAHPHVSEH---   93
S.cerevisiae       FIYKQIGKVSALKEEITKQN-PNRDVTFVSHCGIAHTRWATHGRPEQVNCHPQRSDPEDQ  109
C.albicans         IIVKTTGKVKVLKQKIIDDQ-IDRSAIFDNHVGIAHTRWATHGQPKTENCHPHKSDPKGE  119
Human              QLIKKKGKVKALDEEVHKQQDMDLDIEFDVHLGIAHTRWATHGEPSPVNSHPQRSDKNNE  117
Mouse              QLIKKKGKVKALDEEVHKQQDMDLDIEFDVHLGIAHTRWATHGEPNPVNSHPQRSDKNNE  117
                      :  *::  * :             .    **.******** *   *.**: :

B.subtilisTISTR25  FTLVHNGVIENYVQLKREYLEN-VELKSDTDTEVVVQMIEQFVAG------GLSTEEAFR  146
B.subtilis168      FTLVHNGVIENYVQLKQEYLQD-VELKSDTDTEVVVQVIEQFVNG------GLETEEAFR  146
R.meliloti         VAVVHNGIIENFAELKDELAAGGAEFQTETDTEVVAHLLAKYRRD------GLGRREAMH  146
E.coli             IVVVHNGIIENHEPLREELKARGYTFVSETDTEVIAHLVNWELKQ------GGTLREAVL  147
S.cerevisiae       FVVVHNGIITNFRELKTLLINKGYKFESDTDTECIAKLYLHLYNTNLQNGHDLDFHELTK  169
C.albicans         FIVVHNGIITNYAALRKYLLSKGHVFESETDTECIAKLFKHFYDLNVKAGVFPDLNELTK  179
Human              FIVIHNGIITNYKDLKKFLESKGYDFESETDTETIAKLVKYMYDN--RESQDTSFTTLVE  175
Mouse              FIVIHNGIITNYKDLKKFLESKGYDFESETDTETIAKLVKYMYDN--WESQDVSFTTLVE  175
                   . ::***:* *.  *:         : ::**** :.::

B.subtilisTISTR25  KTLTLLKGSYAIALFDGENTDTIYVAKNKSPLLIGLGDTFN-------------------  187
B.subtilis168      KTLTLLKGSYAIALFDNDNRETIFVAKNKSPLLVGLGDTFN-------------------  187
R.meliloti         AMLKRVKGAYALAVLFEDDPSTIMAARTG-PLAIGHGNGEM-------------------  186
E.coli             RAIPQLRGAYGTVIMDSRHPDTLLAARSGSPLVIGLGMGEN-------------------  188
S.cerevisiae       LVLLELEGSYGLLCKSCHYPNEVIATRKGSPLLIGVKSEKKLKVDFVDVEFPEENAGQPE  229
C.albicans         QVLHELEGSYGLLVKSYHYPGEVCGTRKGSPLLVGVKTDKKLKVDFVDVEFEAQQQHRPQ  239
Human              RVIQQLEGAFALVFKSVHFPGQAVGTRRGSPLLIGVRSEHKLSTDHIPILY------R-T  229
Mouse              RVIQQLEGAFALVFKSVHFPGQAVGTRRGSPLLIGVRSEHKLSTDHIPILY------R-T  229
                     :  :.*::.             ::   ** :*

B.subtilisTISTR25  ----------------------------------------------------V------  188
B.subtilis168      ----------------------------------------------------V------  188
R.meliloti         ----------------------------------------------------F------  187
E.coli             ----------------------------------------------------F------  189
S.cerevisiae       IPLKSNNKSFGLGPKKAREFEAGSQNAN-------LLPIAANEFNLRHSQSRAFLSEDGS  282
C.albicans         QPQINHN-----GATSAAELG--------------FIPVAPGEQNLRTSQSRAFLSEDDL  280
Human              ------------GKDKKGSCN--------------LSRVD--S-----T-TCLFPVEEK-  253
Mouse              ------------GKDKKGSCG--------------LSRVD--S-----T-TCLFPVEEK-  253
                                                                   .
B.subtilisTISTR25  -------VASDAMAMLQVTNEYVELLDKEMVIVTKDEAVIKN----LDGEVMTRASYIAE  237
B.subtilis168      -------VASDAMAMLQVTNEYVELMDKEMVIVTDDQVVIKN----LDGDVITRASYIAE  237
R.meliloti         -------LGSDAIALAPFTNEITYLIDGDWAVIGKTGVHIFD----FDGNVVERPRQIST  236
E.coli             -------IASDQLALLPVTRRFIFLEEGDIAEITRRSVNIFD----KTGAEVKRQDIESN  238
S.cerevisiae       PTPVEFFVSSDAASVVKHTKKVLFLEDDDLAHIYDGELHIHR-SRREVGASMTRSIQTLE  341
C.albicans         PMPVEFFLSSDPASVVQHTKKVLFLEDDDIAHIYDGELRIHRASTKSAGESTVRPIQTLE  340
Human              --AVEYYFASDASAVIEHTNRVIFLEDDDVAAVVDGRLSIHR-IKRTAGDHPGRAVQTLQ  310
Mouse              --AVEYYFASDASAVIEHTNRVIFLEDDDVAAVVDGRLSIHR-IKRTAGDHPGRAVQTLQ  310
                         ..**  ::   *..   * : : . :      *        *    *

B.subtilisTISTR25  LDASDIEKGTYPHYMLKETDEQPLVMRKIIQTYQD--ENGRLAVAGDVADA---VAEADR  292
B.subtilis168      LDASDIEKGTYPHYMLKETDEQPVVMRKIIQTYQD--ENGKLSVPGDIAAA---VAEADR  292
R.meliloti         AAAFLVDKGNHRHFMEKEIYEQPEVIAIALGHYVN---VIDKSCRSDSDAID--FAGVES  291
E.coli             LQYDAGDKGIYRHYMQKEIYEQPNAIKNTLTGRISH-GQVDLSELGPNADEL--LSKVEH  295
S.cerevisiae       MELAQIMKGPYDHFMQKEIYEQPESTFNTMRGRID--YENNKVILGGLKAWLPVVRRARR  399
C.albicans         MELNEIMKGPYKHFMQKEIFEQPDSAFNTMRGRID--FENCVVTLGGLKSWLSTIRRCRR  398
Human              MELQQIMKGNFSSFMQKEIFEQPESVVNTMRGRVN--FDDYTVNLGGLKDHIKEIQRCRR  368
Mouse              MELQQIMKGNFSSFMQKEIFEQPESVVNTMRGRVN--FDDYTVNLGGLKDHIKEIQRCRR  368
                         ** .  :* **  ***      :    .          .       .

B.subtilisTISTR25  IYIVACGTSYHAGLVGKQYIEMWANVPVEVHVASEFSYNMPLLSKKPLFIFLSQSGETAD  352
B.subtilis168      IYIIGCGTSYHAGLVGKQYIEMWANVPVEVHVASEFSYNMPLLSKKPLFIFLSQSGETAD  352
R.meliloti         LAISCCGTAYLAGLIGKYWFERYARLPVEIAVASEFRYREIPLSPQS-ALFISQSGETAD  350
E.coli             IQILACGTSYNSGMVSRYWFESLAGIPCDVEIASEFRYRKSAVRRNSLMITLSQSGETAD  355
S.cerevisiae       LIMIACGTSYHSCLATRAIFEELSDIPVSVELASDFLDRKCPVFRDDVCVFVSHCGETAD  459
C.albicans         IIMIACGTSYHSCLATRSIFEELTEIPVSVELASDFLDRRSPVFRDDTCVFVSQSGETAD  458
Human              LILIACGTSYHAGVATRQVLEELTELPVMVELASDFLDRNTPVFRDDVCFFLSQSGETAD  428
Mouse              LILIACGTSYHAGVATRQVLEELTELPVMVELASDFLDRNTPVFRDDVCFFISQSGETAD  428
                   : :  ***:* : :  :  :*  : :*  : :**:*  .   :  .   . :*:.*****
```

```
B.subtilisTISTR25  SRAVLVQVKALG-HKALTITNVPGSTLSREADYTLLLHAGPEIAVASTKAYTAQIAVLAI  411
B.subtilis168      SRAVLVQVKALG-HKALTITNVPGSTLSREADYTLLLHAGPEIAVASTKAYTAQIAVLAV  411
R.meliloti         TLASLRYCKAHG-LRIGAVVNARESTMARESDAVFPILAGPEIGVARTKAFTCQLAVLAA  409
E.coli             TLAGLRLSKELGYLGSLAICNVPGSSLVRESDLALMTNAGTEIGVASTKAFTTQLTVLLM  415
S.cerevisiae       TMLALNYCLERG-ALTVGIVNSVGSSISRVTHCGVHINAGPEIGVASTKAYTSQYIALVM  518
C.albicans         SILALQYCLERG-ALTVGIVNSVGSSMSRQTHCGVHINAGPEIGVASTKAYTSQYIALVM  517
Human              TLMGLRYCKERG-ALTVGITNTVGSSISRETDCGVHINAGPEIGVASTKAYTSQFVSLVM  587
Mouse              TLMGLRYCKERG-ALTVGITNTVGSSISRETDCGVHINAGPEIGVASTKAYTSQFVSLVM  587
                   :   *      *      : *   *:: * :.  .   **.**.** ***:* *   *

B.subtilisTISTR25  LASVAAERNG---VDIGFDLVKELGIAANAMEALCDQKD-EMEMIAREYLTVSRNAFFIG  467
B.subtilis168      LASVAADKNG---INIGFDLVKELGIAANAMEALCDQKD-EMEMIAREYLTVSRNAFFIG  467
R.meliloti         LRAGAGKARGTISGDEEQALIKSLAEMPAIMGQVLNSIQPEIEVLSRELSNC-RDVLYLG  468
E.coli             LVANVSRLKG-LDASIEHDIVHGLQALPSRIEQML-SQDKRIEALAEDFSDK-HHALFLG  472
S.cerevisiae       FALSLSDDRVS-KIDR-IEIIQGLKLIPGQIKQVLKLEP-RIKKLCATELKDQKSLLLLG  575
C.albicans         FALSLSNDSIS-RKGRHEEIIKGLQKIPEQIKQVLKLEN-KIKDLCNSSLNDQKSLLLLG  575
Human              FALMMCDDRIS-MQERRKEIMLGLKRLPDLIKEVLSMDD-EIQKLA-TELYHQKSVLIMG  544
Mouse              FALMMCDDRIS-MQERRKEIMLGLKRLPDLIKEVLSMDD-EIQKLA-TELYHQKSVLIMG  544
                   :                  ::  *   .  :  :      .:: :.       :  : :*

B.subtilisTISTR25  RGLDYFVCVEGALKLKEISYIQAEGFAGGELKHGTIALIEEGTPVFALATQEHVNLSIRG  527
B.subtilis168      RGLDYFVCVEGALKLKEISYIQAEGFAGGELKHGTIALIEQGTPVFALATQEHVNLSIRG  527
R.meliloti         RGTSFPLAMEGALKLKEISYIQPKSYAAGQLKHGPYALIDENMPVIVIAPHDRFFDKTVT  528
E.coli             RGDQYPIALEGALKLKEISYIHAEAYAAGELKHGPLALIDADMPVIVVAPNNELLEKLKS  532
S.cerevisiae       RGYQFAAALEGALKIKEISYMHSEGVLAGELKHGVLALVDENLPIIAFGTRDSLFPKVVS  635
C.albicans         RGYQFATALEGALKIKEISYMHSEGVLAGELKHGILALVDEDLPIIAFATRDSLFPKVMS  635
Human              RGYHYATCLEGALKIKEITYMHSEGILAGELKHGPLALVDKLMPVIMIIMRDHTYAKCQN  604
Mouse              RGYHYATCLEGALKIKEITYMHSEGILAGELKHGPLALVDKLMPVIMIIMRDHTYAKCQN  604
                   **  :  .:*****:***:*::.:.  .*:****  **::   *:: .  .:    .

B.subtilisTISTR25  NVKEVAARGANTCIIS--------LKGLEDADDRFILPEVNPALRPLVSVVPLQLIAYYA  579
B.subtilis168      NVKEVAARGANTCIIS--------LKGLDDADDRFVLPEVNPALAPLVSVVPLQLIAYYA  579
R.meliloti         NMQEV-ARGGRIILITDEKG---AAASKLDTMHTIVLPEVDEIIAPMIFSLPLQLLAYHT  584
E.coli             NIEEVRARGGQLYVFADQDA----GFVSSDNMHIIEMPHVEEVIAPIFYTVPLQLLAYHV  588
S.cerevisiae       SIEQVTARKGHPIIICNENDEVWAQKSKSIDLQTLEVPQTVDCLQGLINIIPLQLMSYWL  695
C.albicans         AIEQVTARDGRPIVICNEGD---AIISNDKVHTTLEVPETVDCLQGLLNVIPLQLISYWL  692
Human              ALQQVVARQGRPVVICDKED----TETIKNTKRTIKVPHSVDCLQGILSVIPLQLLAFHL  660
Mouse              ALQQVVARQGRPVVICDKED----TETIKNTKRTIKVPHSVDCLQGILSVIPLQLLAFHL  660
                    :::* ** ..  ::                  : :*.    :  :.  :****:::

B.subtilisTISTR25  ALHRGCDVDKPRNLAKSVTVE  600
B.subtilis168      ALHRGCDVDKPRNLAKSVTVE  600
R.meliloti         AVFMGTDVDQPRNLAKSVTVE  605
E.coli             ALIKGTDVDQPRNLAKSVTVE  609
S.cerevisiae       AVNKGIDVDFPRNLAKSVTVE  716
C.albicans         AVNRGIDVDFPRNLAKSVTVE  713
Human              AVLRGYDVDFPRNLAKSVTVE  681
Mouse              AVLRGYDVDFPRNLAKSVTVE  681
                   *:  * *** ***********
```

Figure 6. Alignment of the predicted amino acid sequence of *B. subtilis* TISTR25 *GcaA* open reading frame with those of 7 organisms. The amino acid sequences were aligned by using the ClustalX (1.64b) program. The numbers at the right-hand side are from initial methionine residue for each sequence. Gaps introduced to maximize this alignment are shown by dashes, identical residues are indicated by asterisks, and similar residues are indicated by dots.

Table 2. Estimated genetic distances among nucleotide sequences of the *GcaA* gene from 8 organisms.

	Human	Mouse	*S. cerevisiae*	*C.albicans*	*B. subtilis* TISTR 25	*B. subtilis* 168	*R. meliloti*	*E. coli*
Human	-							
Mouse	0.0972	-						
S. cerevisiae	0.6511	0.6800	-					
C. albican	0.6506	0.6837	0.3666	-				
B. subtilis TISTR 25	1.1735	1.2018	1.1549	1.1987	-			
B. subtilis 168	1.1185	1.1457	1.1091	1.0688	0.1705	-		
R.. meliloti	1.2731	1.2144	1.1958	1.3233	0.9000	0.9302	-	
E. coli	1.0678	1.0592	1.0764	1.1147	1.0080	1.0579	0.8679	-

Table 3. Estimated genetic distances among amino acid sequences of the *GcaA* protein from 8 organisms.

	Human	Mouse	*S. cerevisiae*	*C. albicans*	*B. subtilis* TISTR 25	*B.subtilis* 168	*R. meliloti*	*E. coli*
Human	-							
Mouse	0.01002	-						
S. cerevisiae	0.63101	0.63548	-					
C. albicans	0.63200	0.63106	0.33108	-				
B. subtilis TISTR 25	1.38190	1.38313	1.36734	1.37984	-			
B. subtilis 168	1.38757	1.38449	1.36790	1.36973	0.06940	-		
R.. meliloti	1.32260	1.31277	1.31340	1.34859	1.10600	1.07633	-	
E. coli	1.12711	1.13071	1.15064	1.13109	1.09629	1.08505	0.87441	-

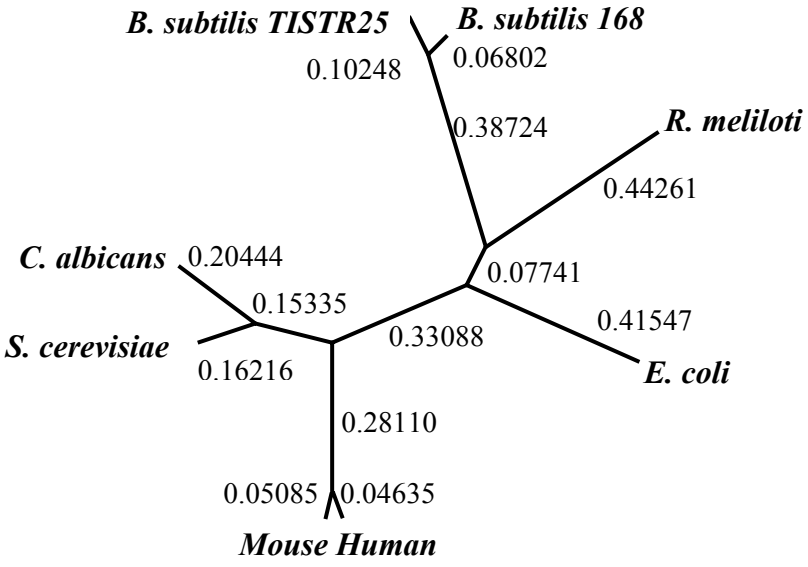


Figure 7. A phylogenetic tree of the *GcaA* gene from *B. subtilis* TISTR 25 and 7 organisms based on nucleotide sequence divergences.

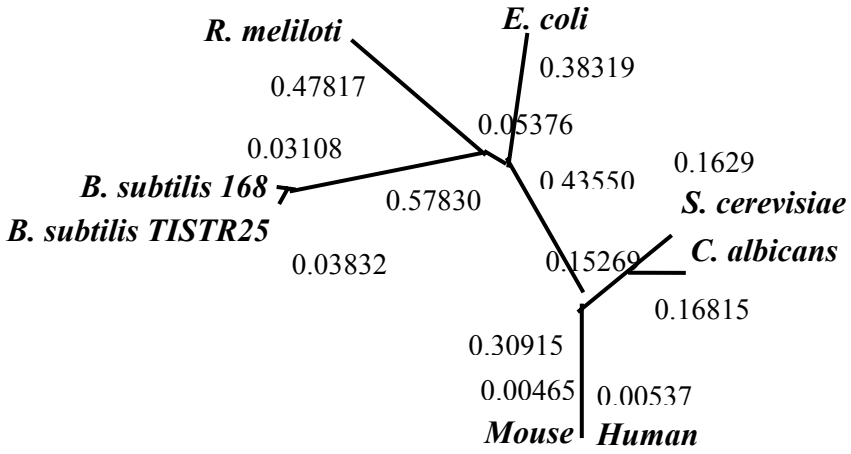
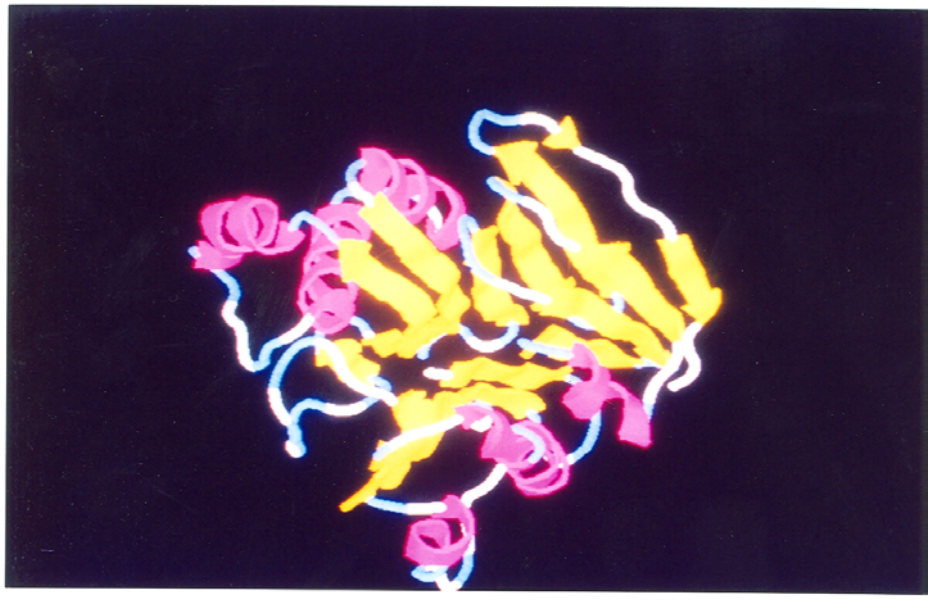
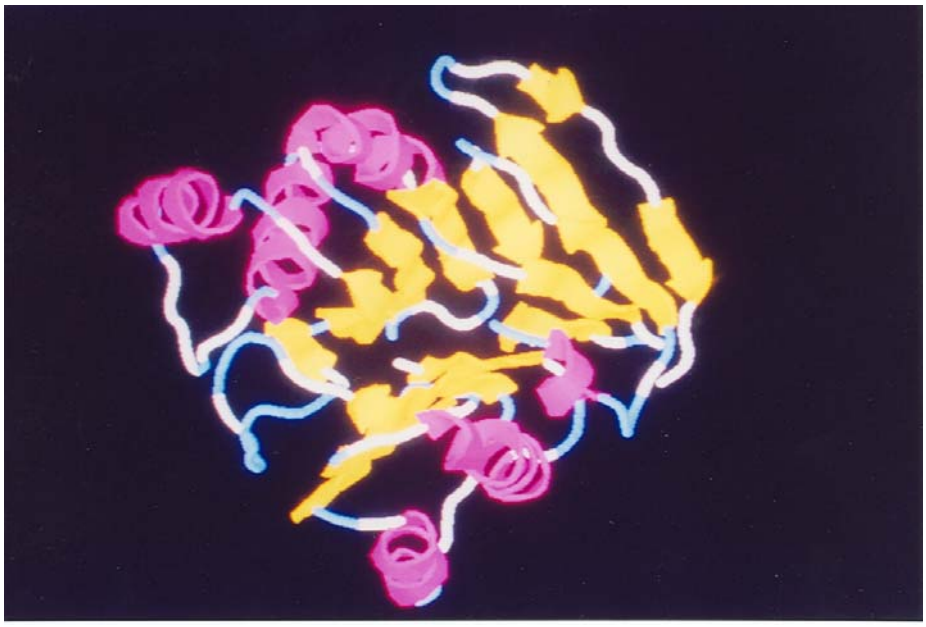


Figure 8. A phylogenetic tree of GcaA enzyme from *B. subtilis* TISTR 25 and 7 organisms based on amino acid sequence divergences.



(a)



(b)

Figure 9. The predicted three-dimensional structure of N-terminal domain of *B. subtilis* TISTR 25 GcaA enzyme (237 amino acids residues) (a) *E. coli* GcaA enzyme in chain A (238 amino acids residues) (b).

L-glutamine D-fructose-6-phosphate amidotransferase activity

The GcaA activities of the clones in various conditions are shown in Table 4. The specific activity of the GcaA enzyme of *E.coli* haboring pUC18 was 96% less than that of pCSBC14. The specific activities of the GcaA enzyme of *E.coli* haboring pCSBC14 with and without IPTG were not significantly different (36.01 and 36.23 units/mg protein).

Table 4. L-glutamine D-fructose-6-phosphate amidotransferase activity in cell extracts.

E. coli DH5α harboring	Total protein (mg)	Total activity (units)	Specific activity (units/mg)
pUC18	23.60	34.25	1.45
pUC18+IPTG	21.60	28.69	1.33
pCSBC14	18.88	679.86	36.01
pCSBC14+IPTG	15.36	556.55	36.23

DISCUSSIONS

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 manually in both directions except for mCSBC143 which was unidirectionally sequenced using the forward primer. The sequence of each clone was compared with the GenBank deposited DNA sequences using BLAST. Each of the sequences was homologous to the L-glutamine D-fructose-6-phosphate amidotransferase (*gcaA*) gene of *B. subtilis* 168 in various positions except the 5' sequence of pCSBC14 which showed no significant similarity to any sequence in the GenBank. Each sequence was combined in the same direction 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, encoding a protein of 600 amino acids (Figure 3) with high homology to GcaA enzyme of *B. subtilis* 168 (86% identity at the nucleotide sequence level). The *B. subtilis* TISTR25 enzyme was also homologous to the *R. meliloti*, *E. coli*, *C. albicans*, *S. cerevisiae*, human and mouse (Figure 6). 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*).

The specific activity of GcaA enzyme from *E. coli* DH5α harboring pUC18 was very low (Table 4). This result was supported by the work of Badet *et al.*[1] who purified the *E. coli* GcaA enzyme. 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 enzyme from *E. coli* DH5α harboring pCSBC14 resulted from the overexpression of this gene from pCSBC14. Since the direction of the cloned *gcaA* gene in pCSBC14 was the same as the *lac* promoter, it was suspected that the expression might be driven by the *lac* promoter. To test this possibility, IPTG was included in the culture medium. The result from Table 4 showed that the enzyme activities of cell extracts from *E. coli* DH5α harboring pCSBC14 with and without IPTG were not significantly 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 to be TTGACT and TTGAAA at -10 and -35, respectively, which were upstreamed from the putative transcription start site G(+1) (Figure 3). 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. A recenty published genome sequence of *B. subtilis* revealed that the *gcaA*

gene (*glmS*) was about 200 kb from the origin of replication (*oriC*).[26]

Comparison of prokaryotic and eukaryotic GcaA enzyme sequences (Figure 6) 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 enzymes differed from the bacterial enzymes in that the former were subjected to allosteric inhibition by uridine 5'-diphosphate *N*-acetylglucosamine (UDP-GlcNAc),[2] it seemed possible that this region contained amino acids residues involved in the interaction with the allosteric effector.

In *E. coli*, it has been reported that methionine1 (at N-terminal) of GcaA enzyme was post translationally removed, leaving cysteine as the N-terminal residue.[1] This cysteine residue was highly conserved among GcaA enzymes. 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 enzyme 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 enzyme was proposed to be involved in the binding of D-fructose-6-phosphate.[27] The Lysine 604 was in the highly conserved C-terminal region. The corresponding residue in *B. subtilis* TISTR25 GcaA enzyme was Lysine 596.

The genetic distances among nucleotide sequences and amino acid sequences of 8 organisms were calculated using the Kimura 2-parameter and the Dayhoff PAM matrix in Phylip version 3.5c, respectively and used to construct a phylogenetic tree. From the phylogenetic tree based on nucleotide sequence divergences (Figure 7), 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 8). This was due to the degeneracy of the amino acid coding sequence. Mutation in DNA may or may not result in amino acid changes. Therefore, mutations in the DNA sequence were reflected in the calculated phylogenetic tree, but the same mutations did not result in changes in amino acid.

The three-dimensional structures of both the *B. subtilis* TISTR 25 and the *E. coli* GcaA enzyme in Figure 9 are shown as an 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 glutamine-6-phosphate.

ACKNOWLEDGMENTS

This work was supported by the Genetic engineering research unit, Department of Biochemistry, Faculty of Science, Chulalongkorn University.

REFERENCES

1. Badet, B., Vermoote, P., Haumont, P.-Y., Lederer, F. and Goffic, F. L. (1987) "Glucosamine synthetase from *Escherichia coli*: purification, properties and glutamine-utilizing site location." *Biochemistry* **26**, 1940-1948.
2. Kornfeld, R. **(**1967) "Studies on L-glutamine D-fructose 6-phosphate amidotransferase." *J. Biol. Chem.* **242**, 3135-3141.
3. Endo, A., Kakiki, K. and Misato, T. (1970) "Feedback inhibition of L-glutamine D-fructose 6-phosphate amidotransferase by uridine diphosphate *N*-acetylglucosamine in Neurospora crassa." *J. Bacteriol.* **103**, 588-594.
4. Crolle, G., D'Este, E. (1980) "Glucosamine sulphate for the management of arthrosis: a controlled clinical investigation." *Curr. Med. Res. Opin.* **7,** 104 -114.
5. D'Ambrosio, E., Casa, B., Bompani, R., Scali, G. and Scali, M. (1981) "Glucosamine sulphate: a controlled clinical investigation in arthrosis." *Pharmatherapeutica* **2**, 504-508.
6. Jung, M. E., Trifunovich, I. D., Gardiner, J. M. and Clevenger, G. L. (1990) "Preparation of modified nucleosides from glucosamine: rapid and efficient formal total synthesis of Several 2'-deoxy *C*-nucleosides." *J. Chem. Soc. Chem. Commun.* 84-85.
7. Walker, J. E., Gay, N. J., Saraste, M. and Eberle, A. N. (1984) "DNA sequence around the *Escherichia coli unc* operon." *Biochem. J.* **224**: 799-815.

8. Baev, N., Endre, G., Petrovics, G., Banfalvi, Z. and Kondorosi, A. (1991) "Six nodulation genes of *nod* box locus 4 in *Rhizobium meliloti* are involved in nodulation signal production: *nodM* codes for D-glucosamine synthetase." *Mol. Gen. Genet.* **228**, 113-124.
9. Watzele, G. and Tanner, W. (1989) "Cloning of the glutamine: fructose-6-phosphate amidotransferase gene from yeast." *J. Biol. Chem.* **264**, 8753-8758.
10. Smith, R. J., Milewski, S., Brown, A. J. P. and Gooday, G. W. (1996) "Isolation and characterization of the *GFA1* Gene encoding the glutamine:fructose-6-phosphate amidotransferase of *Candida albicans*." *J. Bacteriol.* **178**, 2320-2327.
11. McKnight, G. L., Mudri, S. L., Mathewest, S. L., Traxinger, R. R., Marshall, S., Sheppard, P. O. and O'Hara, P. J. (1992) "Molecular cloning, cDNA sequence and bacterial expression of human glutamine: fructose-6-phosphate amidotransferase." *J. Biol. Chem.* **267**, 25208-25212.
12. Sayeski, P. P., Paterson, A .J. and Kudlow, J. E. **(**1994) "The murine glutamine: fructose-6-phosphate amidotransferase-encoding cDNA sequence." *Gene.* **140**, 289-290.
13. Sayeski, P. P., Wang, D., Su, K., Han, I.-O. and Kudlow, J. E. (1997) "Cloning and partial characterization of the mouse glutamine:fructose-6-phosphate amidotransferase (GFAT) gene promoter." *Nucleic Acids Res.* **25**, 1458-1466.
14. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) "Molecular cloning:A laboratory mannual, 2nded." Cold Spring Habor Laboratory Press. New York.
15. Dower, W. J., Miller, J. F. and Ragsdale, C. W. (1988) "High efficiency transformation of *E. coli* by high voltage electroporation." *Nucleic Acids Res.* **16**, 6127-6145.
16. Tanunat, D. (1995) "Cloning of protease gene from *Bacillus subtilis* TISTR25." *Master's Thesis, Department of Biochemistry, Faculty of Science, Chulalongkorn University.*
17. Sanger, F., Nicklen, S. and Coulson, A.R. 1977. "DNA sequencing with chain-terminating inhibitors." *Proc. Natl. Acad. Sci. USA.* **74**, 5463-5467.
18. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) "Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors." *Gene*. **33**, 103-119.
19. Marck, C. (1988) " 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the apple macintosh family of computers." *Nucleic Acids Res.* **16**, 1829-1836.
20. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) "Basic local alignment search tool." *J. Mol. Biol.* **215**, 403-410.
21. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic Acids Res.* **25**, 3389-3402.
22. Higgins, D. G. and Sharp, P. M. (1988) "CLUSTAL: a package for performing multiple sequence alignment on a microcomputer." *Gene*. **73**, 237-244.
23. Ghosh, S., Blumenthal, H. J., Davidson, E. and Roseman, S. (1960) "Glucosamine Metabolism." *J. Biol. Chem.* **235**, 1265-1273.
24. Kenig, M., Vandamme, E. and Abraham, E. P. (1975) "The Mode of Action of Bacilysin and Anticapsin and Biochemical Properties of Bacilysin-resistant Mutants." *J. Gen. Microbiol.* **94,** 46-54.
25. Bradford, M. M. (1976) "A rapid and sensitive method for the quantitation of microgram quantities of protein Utilizing the principle of protein-dye binding." *Anal. Biochem.* **72,** 248-254.
26. Kunst, F. *et al*. **(**1997) "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*." *Nature 390,* 249-256.
27. Golinelli-Pimpaneau, B. and Badet, B. (1991) "Possible involvement of Lys603 from *Escherichia coli* glucosamine-6-phosphate synthase in the binding of its substrate fructose 6-phosphate." *Eur. J. Biochem.* **201**, 175-182.

Received: October 30, 2002
Accepted: September 28, 2004