HAYATI Journal of Biosciences June 2010
Vol. 17 No. 2, p 58-62
EISSN: 2086-4094

Available online at:
http://journal.ipb.ac.id/index.php/hayati
DOI: 10.4308/hjb.17.2.58

Cloning of *araA* Gene Encoding L-Arabinose Isomerase from Marine *Geobacillus stearothermophilus* Isolated from Tanjung Api, Poso, Indonesia

DEWI FITRIANI, BUDI SAKSONO*

Research Center for Biotechnology, Indonesia Institute of Science, Jalan Raya Bogor Km. 46, Cibinong 16911, Indonesia

Received November 30, 2009/Accepted May 5, 2010

L-arabinose isomerase is an enzyme converting D-galactose to D-tagatose. D-tagatose is a potential sweetener-sucrose substitute which has low calorie. This research was to clone and sequence *araA* gene from marine bacterial strain *Geobacillus stearothermophilus* isolated from Tanjung Api Poso Indonesia. The amplified *araA* gene consisted of 1494 bp nucleotides encoding 497 amino acids. DNA alignment analysis showed that the gene had high homology with that of *G. stearothermophilus* T6. The enzyme had optimum activity at high temperature and alkalin condition.

Key words: cloning, *araA* gene, marine bacterium, *Geobacillus stearothermophilus*

INTRODUCTION

L-Arabinose isomerase (L-AI) (EC 5.3.1.4) is an intracelluler enzyme that catalyzes the reversible isomerization of L-arabinose to L-ribulose (Izumori *et al.* 1978). It also can convert D-galactose to D-tagatose (Cheetam & Wootton 1993).

Although the main conversion of L-arabinose isomerase is L-arabinose to L-ribulose, production of D-tagatose is very important as a-low calorie sweeteners. D-tagatose has sucrose-like taste, but does not contribute to calorie production (Levin 2002). In addition, D-tagatose is an anti-hyperglicemial factor with a very low calorie carbohydrate and bulking agent (Levin *et al.* 1995). It is also an efficient antibiofilm which can be used as a cytoprotectants for storage of organs to reduce the referfusion injury (Levin *et al.* 2000). D-tagatose is interested as material for food and drug industry (Rhimi & Bejar 2005).

Thermophilic L-AI has been reported possessing a catalytic activity for conversion of D-galactose to D-tagatose. Generally, isomerization process performed at high temperature (> 70 ^oC) offers several advantages, such as higher conversion yield, faster reaction rate, and lower viscosity of the substrate in the product stream. However, higher-temperature process introduces undesired effect like browning and unwanted by-product formation (Liu *et al.* 1996). In order to overcome these problems a thermostable L-AI with acidic pH optimum (pHopt) would be desirable and crucial for industrial application (Lee *et al.* 2005).

Many researches have been reported the thermophile L-AIs bacteria i.e. L-arabinose isomerase *B. stearothermophilus* US100 (Rhimi 2005), *Geobacillus stearothermophilus* (Jung *et al.* 2005), *G. thermodenitrificans* (Kim *et al.* 2005), *Thermus* sp. (Kim *et al.* 2003b), *Thermoanaerobacter mathranii* (Jorgensen *et al.* 2004) and the acidic L-AI from *Alicyclobacillus acidocaldarius* (Lee *et al.* 2005). Moreover, those of L- AI had been purified and characterized.

L-arabinose isomerase from *G. stearothermophilus* (GSAI) has the highest level of tagatose production and productivity. The production of tagatose is about 230 g/l (Kim *et al.* 2003a) and the productivity is about 54 g/l/h (Ryu *et al.* 2003) using a bioreactor containing immobilized *G. stearothermophilus* L-AI. These results approach commercial production criteria.

The GSAI had been cloned, expressed and characterized. This paper reported the cloning sequencing and amino acid sequence analysis of L-arabinose isomerase from marine bacterium *G. stearothermophilus* isolated from Tanjung Api, Poso, Indonesia.

MATERIALS AND METHODS

Bacterial Strain and Culture Condition. Marine bacterium *G. stearothermophilus* was isolated from Tanjung Api, Poso, Indonesia. It was grown in Nutrient Broth (NB) on an incubator shaker at 55 ^oC. pGEM-T Easy was used as a cloning vector and *Escherichia coli* DH5á was used as a host for cloning. This strain was grown in Luria Bertani (LB) medium on an incubator shaker at 37 ^oC.

Cloning Gene. Genomic DNA of *G. stearothermophilus* was isolated from cells by using Genomic DNA purification Kit (Fermentas) according to the manufacturer's instruction. A modification was performed by adding lysozyme for 5 hours to lyse the cell wall of bacteria. To amplify the *araA* gene, we designed two primers from *G. stearothermophilus* T6 (GeneBank accession number AF160811). The oligonucleotides primers were AraA-F

*Corresponding author. Phone: +62-21-8754587, Fax: +62-21-8754588, E-mail: budisaksono@yahoo.com

(GAACGGGATCCAGCAATGATGCTG) and AraA-R (ATCATACCGCCCCCGCCAAA) with a restriction site of *Bam*H1 in it (underlined). Dreamtaq DNA polymerase (Fermentas) was used to amplify this gene. The PCR conditions were set up as follow; the DNA amplification in 35 cycles, denaturation at 94 ^{o}C for 30s, annealing at 52 ^{o}C for 30s, elongation at 72 ^{o}C for 1 min, and final elongation at 72 ^{o}C for 7 min. The PCR product was purified by using DNA Gel Extraction Kit (Fermentas). The purified PCR product was then inserted into pGEM-T Easy vector and transformed into *E. coli* DH5á. Clones were screened using blue-white selection. The positive clone was showed as white colony in LB plate containing amphicillin 100 mg/ml, Isoprophyl-â-D-thiogalactopyranosidase (IPTG) 0,1 mM and 5-bromo-4-chloro-3-indolyl-â-D-galactopyranoside (X-gal) 2 µg/ml. The clones were checked by colony PCR and plasmid isolation. Further confirmation was done by sequencing of the amplified gene from positive clone.

DNA Sequence Analysis. The amplified DNA gene from positive clone was sequenced at 1stBase Singapore. The sequences were analyzed by multiple sequence alignment using the Clustal W Bioedit 7.0.3.5 program. Amino acid homology of the *araA* gene was determined by using BLASTp (http://www.ncbi.nml.nih.gov/blast).

RESULTS

The PCR product of *araA* gene from the marine strain *G. stearothermophilus* isolated from Tanjung Api, Poso, Indonesia was 1512 bp (Figure 1). The sequence of the gene was shown in Figure 2.

Based on amino acid analysis using BLASTp, the amino acid sequence of *araA* gene of the bacterium exhibited a higher degree of similarity to the *araA* gene of thermophilic bacteria i.e. *G. stearothermophilus* T6 (98%), *Alicyclobacillus acidocaldarius* (97%), *Thermus* sp. (96%), and *Geobacillus thermodenitrificans* (95%) than those of hyperthermophiles i.e. *Thermatoga neopolitana* (63%) and *Thermatoga maritima* (62%) or the mesophiles i.e. *Bacillus halodurans* (68%) and *E. coli* (61%) (Table 1).

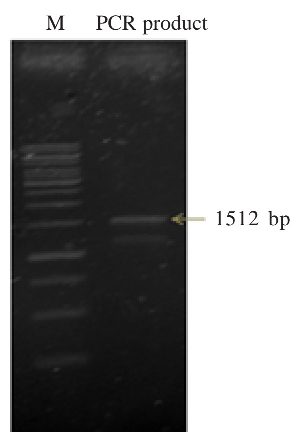

Figure 1. The PCR product of *araA* gene from Indonesian marine bacterial strain *G. stearothermophilus* isolated from Tanjung Api, Poso, Indonesia. Marker 1 kb DNA ladder (lane 1), PCR product (lane 2).

Most of the bacteria having high similarity of their *ara*A gene with the marine strain *G. stearothermophilus* were thermophilic bacteria and active in alkalin condition, only *Alicyclobacillus acidocaldarius* was active in acidic condition (Table 2).

Amino acid sequence analysis using clustalW multiple alignment revealed that 4 amino acids of the marine GSAI different from other thermophiles at the position of R26, S208, K361 and P481. Interestingly, the amino acid of S208 (Serine) only found in this marine GSAI. The catalytic site of the enzyme at E307, E332, H349, H448, D309, E352, and H447 were conserved for L-AIs (Figure 3).

DISCUSSION

This research was important as primary study to screen L-arabinose isomerase as a novel tagatose producer in Indonesia. As a low calorie sweetener, tagatose has potency to substitute sucrose for diabetic and obesity sufferers. The main character of L-AI to convert D-galactose to D-tagatose is at high temperature. The marine bacterial strain *G. stearothermophilus* isolated from Tanjung api, Poso was found in the sea around a mountain. This bacterium lives at high temperature, so that it has potency to produce a thermophile L-AI.

Generally, isomerization is performed at high temperature, so that thermophile L-AI is suitable for this process. Isomerization at high temperature offers several advantages such as higher conversion yield, faster reaction rate, and lower viscosity of the substrate (Liu *et al.* 1996). Previous study of L-AI from *G. stearothermophilus* (GSAI) found that GSAI is suitable for commercial production of D-tagatose because it has high conversion of D-galactose to D-tagatose (Kim *et al.* 2003a; Ryu *et al.* 2003).

This study was succed to clone and sequence *araA* gene from Indonesian marine bacterial strain *G. stearothermophilus*. Based on BLASTp analysis, L-AI from the bacterium has high identity with most thermophile L-AIs which have optimum activity in high temperature and alkalin condition. The sequence analysis found 4 amino acids existed in the L-AI from the bacterium differ from other thermophile L-AIs. The differences of these amino acids were at position of R26, S208, K361, and P481. We found specific amino acid present in our AI at the position of 208, that is Serine residue, whereas others GSAI was Asparagine and Threonine in *B. halodurans.* These amino acids might be related to catalytic activity of L-AI because the position of these amino acids was close to catalytic site of L-AI (Rhimi *et al.* 2007). The three-dimentional (3D) protein structure analysis and enzyme assay may prove the statement.

Rhimi *et al.* (2007) found that the catalytic site of BSAI us100 were at position of E306, E331, H348, and H447. The other amino acids i.e. D308, E351, F329, and H446 also contributed to catalytic activity. This catalytic site was conserved with those of marine local GSAI at position of E307, E332, H349, and H448, as well as at position of D309, E352, F330, and H447.

```
1     ATG ATG CTG TCA TTA CGT CCT TAT GAA TTT TGG TTT GTA ACG GGA   45
1      M   M   L   S   L   R   P   Y   E   F   W   F   V   T   G    15

46    AGC CAG CAC TTG TAC GGA GAA GAA GCA TTA AGG CAA GTT GAA GAG   90
16     S   Q   H   L   Y   G   E   E   A   L   R   Q   V   E   E    30

91    CAT TCA ATG ATG ATT GTC AAT GAG CTG AAT CAA GAT TCA GTG TTC   135
31     H   S   M   M   I   V   N   E   L   N   Q   D   S   V   F    45

136   CCG TTC CCA CTT GTT TTC AAA TCA GTT GTC ACA ACA CCA GAG GAA   180
46     P   F   P   L   V   F   K   S   V   V   T   T   P   E   E    60

181   ATT CGG CGC GTT TGC CTT GAG GCG AAT GCG AGC GAA CAA TGC GCT   225
61     I   R   R   V   C   L   E   A   N   A   S   E   Q   C   A    75

226   GGG GTC ATC ACT TGG ATG CAT ACA TTC TCG CCA GCG AAG ATG TGG   270
76     G   V   I   T   W   M   H   T   F   S   P   A   K   M   W    90

271   ATT GGC GGC CTT TTG GAG CTG CGA AAA CCG TTA TTG CAT CTT CAC   315
91     I   G   G   L   L   E   L   R   K   P   L   L   H   L   H    105

316   ACT CAA TTT AAC CGT GAT ATT CCG TGG GAC AGC ATC GAT ATG GAC   360
106    T   Q   F   N   R   D   I   P   W   D   S   I   D   M   D    120

361   TTT ATG AAC TTA AAC CAA TCG GCT CAC GGT GAC CGG GAA TAC GGA   405
121    F   M   N   L   N   Q   S   A   H   G   D   R   E   Y   G    135

406   TTT ATC GGC GCG AGA ATG GGC GTG GCC CGG AAA GTG GTG GTC GGG   450
136    F   I   G   A   R   M   G   V   A   R   K   V   V   V   G    150

451   CAC TGG GAA GAC CCA GAA GTC CGC GAG CGG CTG GCG AAA TGG ATG   495
151    H   W   E   D   P   E   V   R   E   R   L   A   K   W   M    165

496   CGG ACG GCT GTC GCG TTT GCG GAA AGC CGC AAC CTA AAA GTG GCT   540
166    R   T   A   V   A   F   A   E   S   R   N   L   K   V   A    180

541   CGT TTT GGC GAC AAC ATG CGT GAA GTG GCT GTG ACG GAA GGG GAC   585
181    R   F   G   D   N   M   R   E   V   A   V   T   E   G   D    195

586   AAA GTC GGA GCG CAA ATT CAA TTC GGC TGG TCG GTC AGC GGC TAT   630
196    K   V   G   A   Q   I   Q   F   G   W   S   V   S   G   Y    210

631   GGC ATC GGG GAT TTG GTG CAA TAC ATC CGC GAT GTT TCT GAA CAA   675
211    G   I   G   D   L   V   Q   Y   I   R   D   V   S   E   Q    225

676   AAA GTG AAC GAG TTG CTC GAT GAA TAC GAG GAG CTG TAC GAC ATT   720
226    K   V   N   E   L   L   D   E   Y   E   E   L   Y   D   I    240

721   GTA CCC GCC GGC CGC CAA GAA GGG CCC GTT CGC GAA TCA ATT CGT   765
241    V   P   A   G   R   Q   E   G   P   V   R   E   S   I   R    255

766   GAA CAG GCG CGG ATT GAA CTC GGG CTG AAA GCC TTT TTG CAG GAT   810
256    E   Q   A   R   I   E   L   G   L   K   A   F   L   Q   D    270

811   GGG AAC TTT ACC GCT TTC ACG ACG ACG TTC GAG GAC TTG CAT GGG   855
271    G   N   F   T   A   F   T   T   T   F   E   D   L   H   G    285

856   ATG AAG CAG CTC CCG GGA CTT GCC GTT CAG CGA CTC ATG GCG GAA   900
286    M   K   Q   L   P   G   L   A   V   Q   R   L   M   A   E    300

901   GGC TAC GGC TTT GGC GGC GAA GGC GAC TGG AAA ACA GCC GCC CTC   945
301    G   Y   G   F   G   G   E   G   D   W   K   T   A   A   L    315

946   GTC CGG TTG ATG AAA GTC ATG GCC GAC GGC AAA GGG ACG TCG TTC   990
316    V   R   L   M   K   V   M   A   D   G   K   G   T   S   F    330

991   ATG GAA GAT TAC ACG TAT CAC TTC GAG CCG GGC AAC GAA CTG ATT   1035
331    M   E   D   Y   T   Y   H   F   E   P   G   N   E   L   I    345

1036  CTC GGC GCT CAT ATG CTC GAA GTA TGT CCG ACG ATC GCG GCA ACC   1080
346    L   G   A   H   M   L   E   V   C   P   T   I   A   A   T    360

1081  AAA CCA AGA ATC GAA GTT CAT CCG CTT TCC ATC GGC GGA AAA GAA   1125
361    K   P   R   I   E   V   H   P   L   S   I   G   G   K   E    375

1126  GAT CCG GCC CGT CTT GTG TTT GAC GGC GGC GAG GGT GCG GCG GTC   1170
376    D   P   A   R   L   V   F   D   G   G   E   G   A   A   V    390

1171  AAC GCG TCA TTG ATC GAC TTA GGG CAC CGT TTC CGA CTC ATC GTC   1215
391    N   A   S   L   I   D   L   G   H   R   F   R   L   I   V    405

1216  AAT GAA GTC GAT GCG GTG AAA CCG GAA CAC GAC ATG CCG AAA TTA   1260
406    N   E   V   D   A   V   K   P   E   H   D   M   P   K   L    420

1261  CCA GTC GCC CGC ATT TTA TGG AAG CCT CGC CCG TCG CTC CGC GAC   1305
421    P   V   A   R   I   L   W   K   P   R   P   S   L   R   D    435

1306  TCC GCT GAA GCA TGG ATT TTA GCT GGC GGC GCC CAC CAT ACG TGC   1350
436    S   A   E   A   W   I   L   A   G   G   A   H   H   T   C    450

1351  TTC TCA TTT GCG GTT ACA ACA GAA CAG CTG CAA GAC TTT GCG GAA   1395
451    F   S   F   A   V   T   T   E   Q   L   Q   D   F   A   E    465

1396  ATG GCA GGG ATT GAA TGT GTC GTG ATC AAT GAA CAT ACG TCC GTC   1440
466    M   A   G   I   E   C   V   V   I   N   E   H   T   S   V    480

1441  CCC TCA TTC AAG AAC GAA CTA AGA TGG AAT GAA GTA TTT TGG CGG   1485
481    P   S   F   K   N   E   L   R   W   N   E   V   F   W   R    495

1486  GGG CGG TAA   1494
496    G   R   …
```

Figure 2. DNA and amino acid Sequences of *araA* gene of Indonesian marine bacterial strain *G. stearothermophilus* isolated from Tanjung Api, Poso, Indonesia.

For industrial application, character of L-AI which was active at high temperature and acidic pH could allow to reduce browning process and formation of by-product (Liu *et al.* 1996). To fulfill the requirements, some studies of genetic engineering and site-directed mutagenesis were performed to improve catalytic activity and enzyme stability at low pH (acidic condition). A study of L-AI from *Alicyclobacillus acidocaldarius* found that amino acid at K269 position played an important role on pH optimum (Lee *et al*. 2005). Rhimi *et al.* (2009) proved this theory by altering the Q268K of the enzyme of BSAI US100 could reduce the optimum pH to be 6.5. The further research

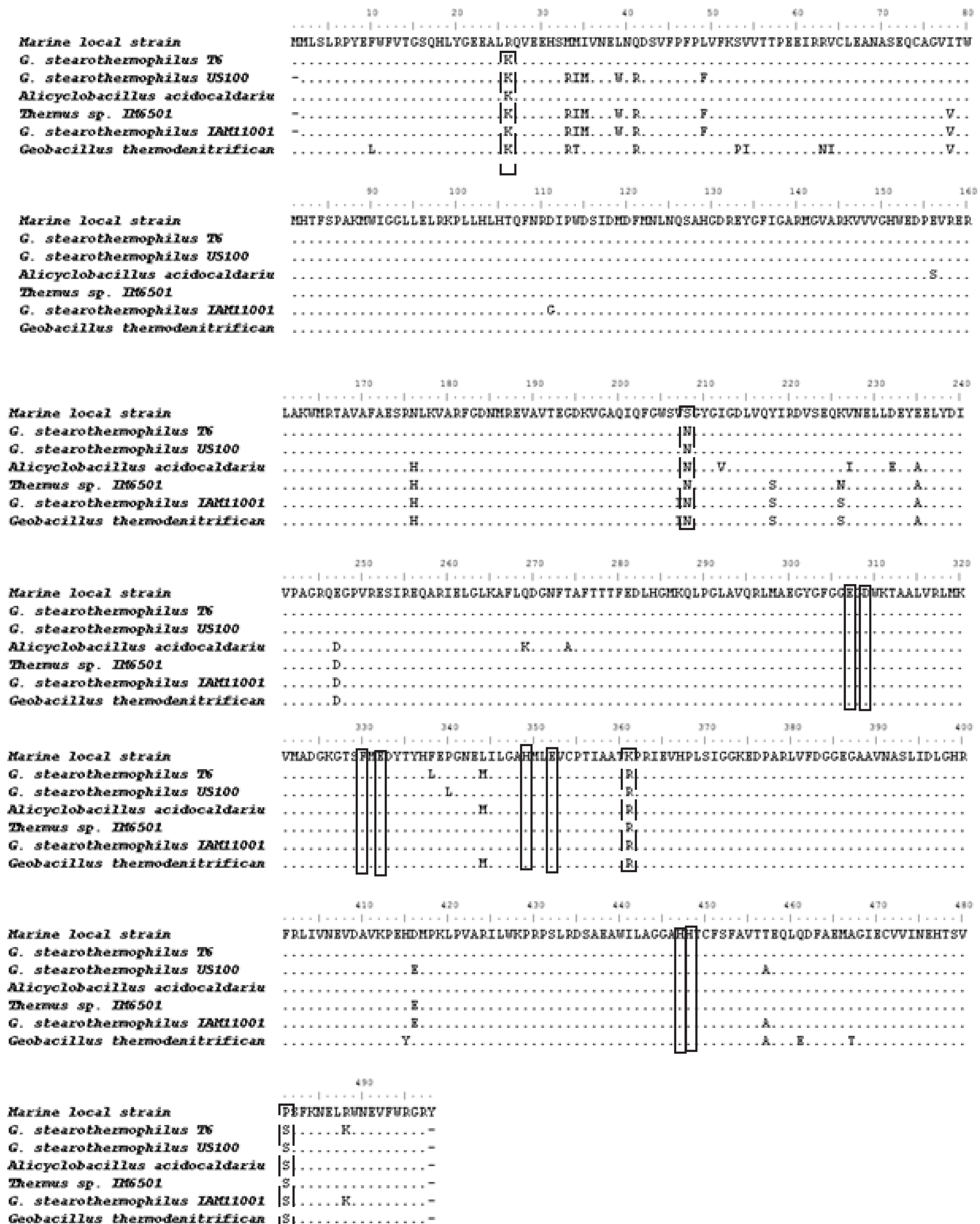

Figure 3. Alignment of amino acid sequences of Indonesia marine bacterium L-arabinose isomerase from *G. stearothermophilus* (GSAI) with other L- arabinose isomerase (L-AIs). The dot box showed the different region of L-AIs. Line box showed the conserved catalytic site of L-AIs.

Table 1. Similarity value of L- arabinose isomerase (L-AIs) of Indonesia marine bacterium L-arabinose isomerase from *G. stearothermophilus* (GSAI) with other bacteria

Character of enzym	Genebank Accesion number	Strain	Similarity (identity) (%)	E value
Thermophile	AAD45718	*Geobacillus stearothermophilus* T6	98	0.0
	AAY68209	*Alicyclobacillus acidocaldarius*	97	0.0
	CAI29261	*Geobacillus stearothermophilus US100*	97	0.0
	AAO72082	*Thermus* sp.	96	0.0
	ABY84698	*Geobacillus stearothermophilus* IAM11001	95	0.0
	AAQ72737	*Geobacillus thermodenitrificans*	95	0.0
Hyperthermophile	AAK18729	*Thermatoga neopolitana*	63	0.0
	NP 228089	*Thermatoga maritime*	62	0.0
Mesophile	NP 242739	*Bacillus halodurans*	68	0.0
	NP 414604	*Escherichia coli*	61	0.0

Table 2. Comparative biochemical characteristics of L-arabinose isomerase from several bacteria

Strain	Optimum temperature (°C)	optimum pH	Requirement of metal ion	Reference
Bacillus halodurans	50	7.5-8.0	Mn^{2+}	Lee *et al*. 2005
Escherichia coli	30	8.0	Fe^{2+}, Mn^{2+}	Yoon *et al*. 2003
Alicyclobacillus acidocaldarius	65	6.0	Mn^{2+}	Lee *et al*. 2005
G. stearothermophilus T6	70	7.0-7.5	Mn^{2+}	Lee *et al*. 2005
G.stearothermophilus US100	80	7.5-8.0	No requirement	Rhimi *et al*. 2005
G. stearothermophilus IAM11001	65	7.5	Mn^{2+}	Kim *et al*. 2003a
G. thermodenitrificans	70	8.5	Mn^{2+}	Kim *et al*. 2005
Thermus sp.	60	8.5	Mn^{2+}	Kim *et al*. 2003b
Thermatoga neapolitana	85	7.0	Co^{2+}	Kim *et al*. 2005
Thermatoga maritime	90	7.0-7.5	Co^{2+}	Lee *et al*. 2005

is to study expression, purification and characterization of marine bacterium GSAI as a potential novel enzyme for tagatose production in Indonesia.

ACKNOWLEDGEMENT

We are greatful to Eko Chasanah from Research Center for Marine and Fisheries Product Processing and Biotechnology, Ministry of Marine Affairs and Fisheries Republic of Indonesia to providing the Indonesian marine bacterial strain *Geobacillus stearothermophilus*.

REFERENCES

Cheetam PS, Wootton AN. 1993. Bioconversion of D-galactose to D-tagatose. *Enzyme MicrobTechnol* 15:105-108.

Izumori K, Ueda YN, Yamanaka K. 1978. Pentose metabolism in *Mycobacterium smegmetis*: comparison of L-arabinose isomerase induced by L-arabinose and D-galactose. *J Bacteriol* 133:413-414.

Jung ES, Kim HJ, Oh DK. 2005. Tagatose production by immobilized recombinant Escherichia coli cells containing Geobacillus *stearothermophilus* L-arabinose isomerase mutant in a packed-bed bioreactor. *Biotechnol Prog* 21:1335-1340.

Jørgensen F, Hancen OC, Stougaard P. 2004. Enzymatic conversion of D-galactose to D-tagatose: heterologous expression and characterization of Thermostable L-arabinose isomerase from *Thermoanaerobacter mathranii*. *Appl Microbiol Biotechnol* 64:816-822.

Kim HJ, Oh DK. 2005. Purification and characterization of an L-arabinose isomerasefrom an isolated strain of *Geobacillus thermodenitrificans* producing D-tagatose. *J Biotechnol* 120:162-173.

Kim HJ, Ryu SA, Kim P, Oh DK. 2003a. A feasible enzymatic process for D-tagatose production by an immobilized thermostable L-arabinose isomerase in a packed-bed bioreactor. *Biotechnol Prog* 19:400-404.

Kim JW, Kim YW, Roh HJ, Kim HY, Cha JH, Park KH, Park CS. 2003b. Production of tagatose by a recombinant thermostable L-arabinose isomerase from *Thermus* sp. IM6501. *Biotechnol Lett* 25:963-967.

Lee SJ, Lee DW, Choe EA, Hong YH, Kim SB, Kim BC, Pyun YR. 2005. Characterization of a thermoacidophilic L-arabinose isomerase from *Alicyclobacillus acidocaldarius*: Role of Lys-269 in pH optimum. *Appl Environ Microbiol* 71:7888-7896.

Levin GV. 2000. Increased fertility and improved fetal development. *US patent* 6225452.

Levin GV. 2002. Tagatose, the new GRAS sweeteners and helath product. *J Med Food* 5:23-36.

Levin GV, Zehner LR, Saunders JP, Beadle JR. 1995. Sugar substitutes: their energy values, bulk characteristics, and potencial health benefits. *Am J Clin Nutr* 126:1601-1609.

Liu S, Wiegel J, Gherardini FC. 1996. Purification and cloning of a thermostable xylose (glucose) isomerization with an acidic pH optimum from *Thermoanaerobacterium* strain JW/SL-YS 489. *J Bacteriol* 178:5938-5945.

Rhimi M, Aghajari N, Juy M, Chouayekh H, Maguin E, Haser R, Bejar S. 2009. Rational design of *Bacillus stearothermophilus* US100 L-arabinose isomerase: Potential applications for D-tagatose production. *Biochimie* 91:650-653.

Rhimi M, Bejar S. 2005. Cloning, purification and biochemical characterization of metallic-ions independent and thermoactive L-arabinose isomerase from the *Bacillus stearothermophilus* US100 strain. *Biochim Biophys Acta* 1760:191-199.

Rhimi M, Juy M, Aghajari N, Haser R, Bejar S. 2007. Probing the essensial catalytic residue and substrate affinity in the thermoactive *Bacillus stearothermophilus* US100 L-arabinose isomerase by site-directed mutagenesis. *J Bacteriol* 189:3556-3563.

Ryu SA, Kim CS, Kim HJ, Baek DH, Oh DK. 2003. Continous D-tagatose production by immobilized thermostable L-arabinose isomerase in a packed-bed bioreactor. *Biotechnol prog* 19:1643-1647.

Yoon SH, Kim P, Oh DK. 2003. Properties of L-arabinose isomerase from Escherichia coli as biocatalysis for tagatose production. *World J Microbial Biotechnol* 19:34-34.