August 2021
Master Dissertation

Purification and preliminary structural studies of hybrid L-arabinose isomerases

Graduate School of Chosun University

Department of Biomedical Sciences

HOANG NGOC KIM QUYEN

Purification and preliminary structural studies of hybrid L-arabinose isomerases

융합형 L-아라비노스 이성화효소의

분리동정 및 초기 구조연구

27th August 2021

Graduate School of Chosun University

Department of Biomedical Sciences

HOANG NGOC KIM QUYEN

Purification and preliminary structural studies of hybrid L-arabinose isomerases

Advisor: Prof. Sung Haeng Lee

This dissertation is submitted to the Graduate School of Chosun University in partial fulfillment of the requirements for the degree of Master in Science.

April 2021

Graduate School of Chosun University

Department of Biomedical Sciences

HOANG NGOC KIM QUYEN

This is to certify that the Master dissertation of Hoang Ngoc Kim Quyen has successfully met the dissertation requirements of Chosun University

Chosun University: Prof. Tae-Hyoung Kim

Chosun University: Prof. Sung Haeng Lee

Chosun University: Prof. Jung Hee Park

May 2021

Graduate School of Chosun University

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

L-AI	L-Arabinose Isomerase
E. coli	*Escherichia coli*
ECAI	*Escherichia coli* L-Arabinose Isomerase
GKAI	*Geobacillus kaustophilius* L-Arabinose Isomerase
AAAI	*Alicyclobacillus acidocaldarius* L-Arabinose Isomerase
ASAI	*Alicyclobacillus sp*. TP-7 L-Arabinose Isomerase
TMAI	*Thermotoga maritima* L-Arabinose Isomerase
LB	Luria-Bertani broth
PMSF	Phenylmethyl sulfonyl fluoride
DTT	Dithiothreitol
Ni-NTA	Nickel-nitrilotriacetic acid
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
IPTG	Isopropyl β-D-1-thiogalactopyranoside

국문 초록

융합형 L-아라비노스 이성화효소의 분리동정 및 초기 구조연구

호앙 응옥 킴 꾸옌
지도교수: 이 성 행, Ph.D.
의과학과
조선대학교 대학원

L-arabinose 이성화 효소 (L-AI) (EC 5.3.1.4)는 생체 외에서 가역적으로 L-arabinose를 L-ribulose로, 생체 내에서 D-galactose를 D-tagatose로 전환시키는 세포 내 효소이다. D-galactose의 천연 keto-hexose 이성질체인 D-tagatose는 현재 의학 및 다이어트 산업에서 저칼로리 감미료로 각광을 받고 있다. 따라서, 최근에는 D-galactose의 기질 특이성, 열 안정성 문제 및 낮은 pH 조건을 극복하고 D-tagatose를 얻고자 L-arabinose 이성화 효소를 개량하는 연구가 집중적으로 이루어지고 있다.

우리는 D-tagatose 생산에 최적의 조건을 가진 L-arabinose 이성화 효소를 찾고자 여러 물리-화학적 특성을 가진 AI들을 디자인하고, 각 L-AI의 물리-화학적 구조-기능 관계를 확인했다.

그 결과 4가지 타입의 모체 L-AI 효소 유전자들 (중온성 *Escherichia coli* AI (ECAI), 호열성 *Geobacillus kaustophilius* AI (GKAI), 호열산성 *Alicyclobacillus acidocaldarius* AI (AAAI)와 열알칼리성 *Alicyclobacillus sp.* TP-

7 AI (ASAI))을 재조합하여 두 종의 융합형 AI (AI-8 및 AI-10) 를 얻었다. 이들 도메인의 물리-화학적인 결합과 구조-기능 관계성을 밝히고자 결정화를 시도했다.

X-선 회절을 통해 얻은 데이터 결과는 두 융합형 L-AI 결정이 모체 L-AI와 동일한 6량체 L-arabinose 이성화 효소 결합체이며, 효소로서 유효한 기능을 할 것으로 예상 했다. 또한, 활성 자리에서 Mn으로 예상되는 전이금속의 전자밀도지도를 확인했고, 이는 기존 당전환 효소에서도 발견되는 결과이다. 더 나아가, arabinose의 아나로그인 adonitol을 soaking한 결정의 활성자리에서는 전이금속이외에 더 큰 전자밀도지도가 발견되었다.

따라서 이 결과를 종합하면, 융합형 L-AI들은 기존 모체 L-AI와 구조적 동일하고, galactose와 구조적으로 결합하여 tagatose를 생산하는 활성을 가질 것으로 사료된다.

뿐만 아니라 본 초기 구조연구에서 얻어진 X-선 회절 데이터를 통해 어떤 융합형 L-AI의 구조적 특징이 D-galactose의 기질 특이성, 열안정성 문제 및 낮은 pH 조건를 극복하고 효소로서 기능을 가능하도록 돕는지 설명할 수 있을 것이다. 더 나아가 이 L-AI들의 도메인이 가지는 특이성을 밝힘으로써, 원하는 물리-화학적 기능을 가진 융합형 L-AI 개발 전망과 산업적 응용을 통한 tagatose을 대량 생산으로 이어질 수 있다.

ABSTRACT

Purification and preliminary structural studies of hybrid L-arabinose isomerases

Hoang Ngoc Kim Quyen

Advisor: Prof. Sung Haeng Lee, Ph.D.

Department of Biomedical Sciences

Graduate School of Chosun University

L-arabinose isomerase (L-AI) (EC 5.3.1.4) is an intracellular enzyme that catalyzes the reversible isomerization of L-arabinose to L-ribulose *in vitro* and conversion of D-galactose to D-tagatose *in vivo* that is a low-calorie bulk sweetener with some remarkable health benefits. Therefore, the enzymatic production of D-tagatose from the low-cost D-galactose has drawn industrial attention using L-arabinose isomerase in recent years. However, the current commercial production of D-tagatose is costly than other sugars despite the health benefit, which is mainly due to the difficulties in optimizing the fermentation condition such as pH, temperature, and substrate specificity from various distinctive L-AIs. One of the ways to solve this problem is to know the relationship between structural domain and functional characteristics governing the physico-chemical properties among L-AIs, followed by the creation of hybrid L-AIs harboring optimized combinations for the industrial application. To meet this purpose, molecular breeding was applied to generate recombined genes, resulting in progenitors with combined properties from parental sources. Among them, two hybrid constructs (hybrid AI-8 and AI-10) having different temperature and pH dependences from their parental enzymes were chosen for the structure study to investigate the structure-function relationship for L-AI

engineering.

Structure analysis of these hybrids can reveal the proposed adaptive traits that contribute to the temperature and pH dependence of L-AIs. In this report, I purified, crystallized, and collected the preliminary X-ray diffraction data of hybrid AI-8 and AI-10. The apo form AI-8 diffracted to 2.5 Å, and belonged to space group C2, with the unit-cell dimensions a = 204.59, b = 81.91, c = 192 Å, α = γ = 90, β = 117.9°, and the calculated Matthews coefficient of 2.24 $Å^3$ Da^{-1} with a solvent content of 45%, indicating six molecules per asymmetric unit. The adonitol-bound AI-8 crystal diffracted to 2.6 Å and belonged to space group C2 with the unit cell parameters a = 205.58, b = 82.12, c = 192.73 Å, α = γ = 90, β = 117.93°, and the Matthews coefficient parameter of 2.26 $Å^3$ Da^{-1} corresponded to 46% solvent content and the presence of six molecules per asymmetric unit. The apo form of hybrid AI-10 crystal diffracted to 3.2 Å resolution, and belonged to space group $C222_1$ with the unit cell dimensions a = 148.48, b = 258.58, c = 165.73 Å, α = β = γ = 90°, and the Matthews coefficient parameter of 2.5 $Å^3$ Da^{-1} indicated the asymmetric unit of six molecules and a calculated solvent content of 51%. The preliminary structure was solved by molecular replacement using apo-GKAI (PDB code 4R1P) and apo-ECAI (PDB code 2AJT) as search models.

The results from this study manifest that the hybrid L-AIs conserve the structural integrity of known L-AIs with novel physico-chemical properties that have not been observed. Therefore, the complete studies would help to elucidate the domain-specific function for the properties and designing novel L-AIs for industrial uses that overcome the current huddle of tagatose production by lowering the cost.

I. INTRODUCTION

An L-arabinose isomerase (L-AI) (EC 5.3.1.4) belongs to the family of isomerases, catalyzing the reversible isomerization of L-arabinose to L-ribulose (Izumori *et al*., 1978) and conversion of D-galactose to D-tagatose (Cheetham & Wootton, 1993, Manjasetty & Chance, 2006, Kim *et al*., 2002) shown in Figure 1.

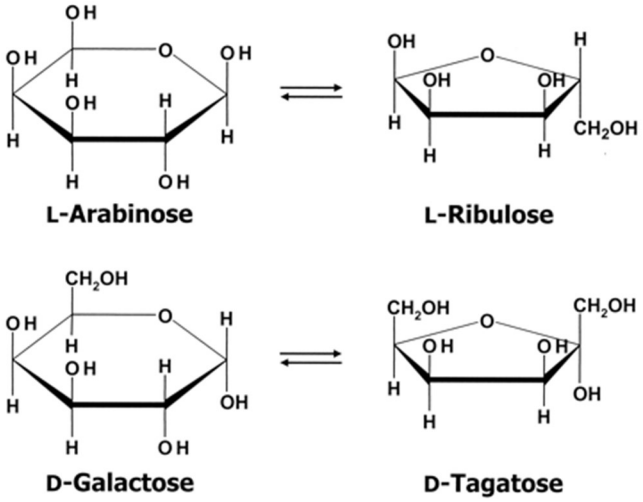


Figure 1: Isomerization of an aldose to a ketose (Manjasetty *et al.*, 2006).

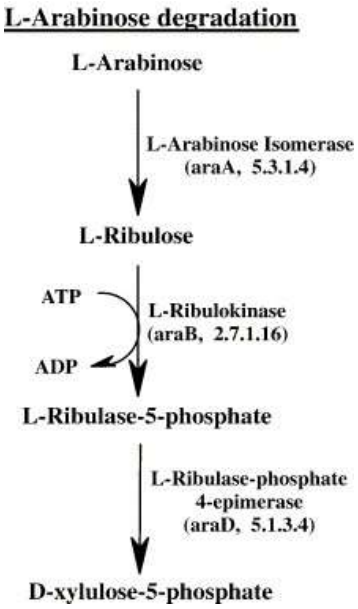


Figure 2: Pentose phosphate pathway (Manjasetty *et al.*, 2006).

In the aldopentose metabolism or pentose phosphate pathway (Figure 2), bacteria utilize arabinose as a carbon source. In the first intracellular step, L-arabinose isomerase catalyzes the in vivo reversible isomerization of L-arabinose into L-ribulose and can be used *in vitro* for the production of D-tagatose using D-galactose as a substrate with the virtue of sharing the similar configurations at the positions of C3 and C4 of L-*cis*-hydroxyl group (Cheetham & Wootton, 1993; Roh *et al.*, 2000). Bacteria utilize arabinose as a carbon source that relates to arabinose operon. Over the past decades of study on the L-arabinose operon of *Escherichia coli* has provided novel insights into the mechanism of positively regulated gene activity. Accordingly, arabinose operon components commonly include some genes: structural genes (araB, araA, and araD) that encode the metabolic enzymes that break down arabinose and regulatory genes (araI, araC, araO) (William *et al*., 2008). In this study, we focus on improving L-AI's specific activity of encoded by gene araA for the commercial D-tagatose production (James W.Patrick, 1968, Sá-Nogueira & Lencastre, 1989, Schlief, 2010, 2000).

D-Tagatose, a natural keto-hexose isomer of D-galactose, is currently being introduced as a low-calorie bulk sweetener (Hichem *et al.*, 2007). The sweetness of this alternative is equivalent to sucrose but the caloric value is around 30% of the energy content of sucrose (Levin *et al*., 2002) and has some considerable health benefits. Because of its low intestinal absorption, it does not facilitate tooth decay or cause an increase in blood glucose levels (Hichem *et al.*, 2007). The above reasons enable D-tagatose to have a low-calorific value for humans and to be approved as a "Generally recognized as safe (GRAS)" substance for humans on consumption, under United States Food and Drug Administration (FDA) regulations (Levin *et al.,* 2002); currently, D-Tagatose is in the market for use as a low-calorie sweetener in confectionery, beverages, health foods, and dietary products in Korea, New Zealand, Australia, and the United States of America. This alternative is helpful, especially for reducing symptoms associated with type II diabetes and combating obesity (Levin, 2002; Kim, 2004; Manjasetty & Chance, 2006). For these reasons, D-tagatose, a prebiotic compound, has gained commercial interest.

On the other hand, the extraction of D-tagatose is not possible due to a lack of available biological sources. The first industrial tagatose production was patented by Beadle *et al*. (1992), which contained a calcium catalyst used to isomerize D-galactose derived from lactose hydrolysis chemically. Additionally, this method has several disadvantages, such as complicated purification steps, low yields, the formation of chemical waste, and by-product release despite its low cost. Such drawbacks have been mostly overcome. In several studies, the effectiveness of organocatalysis in carbohydrate preparation, such as tagatose, has been illustrated. Through this practical and environment-friend organocatalytic strategy, good yields of polyols synthesis and high levels of stereoselectivity have been achieved (Suri *et al*., 2005, 2006). Many research groups have attempted to evolve alternative enzymatic procedures for D-tagatose manufacture. An enzymatic isomerization typically has several advantages, including lower pH and high temperature, no creation of an unexpected product, and particularly the potential for catalysts improvement through molecular evolution. In recent years, D-tagatose is made enzymatically from D-galactose by L-AI. Several L-AI-containing

bacteria have been used for D-tagatose manufacture, including *Escherichia coli*, *Thermotoga neapolitana*, *Geobacillus stearothermophilus*, *Thermus sp*., *Thermoanaerobacter mathranii*, *Thermotoga maritima*, *Geobacillus thermodenitrificans*, *Alicyclobacillus acidocaldarius*, and *Bacillus stearothermophilus* US100 (Roh *et al*., 2000; Kim *et al*., 2002, 2003a, b; Jorgensen *et al*., 2004; Lee *et al*., 2004, 2005a; Kim & Oh, 2005; Rhimi & Bejar, 2006).

Furthermore, recent biotechnological techniques have facilitated the efficient isolation of enzymes with improved catalytic properties. A few characteristics of this enzyme will need to be changed to meet L-AI's industrial demands for tagatose development. The L-AI's specificity for galactose recognition and enzyme thermostability should be enhanced, whereas the pH of the isomerization process should be decreased. Since the galactose-tagatose equilibrium moves toward tagatose at higher temperatures, industrial tagatose processes are recommended to be performed at 60–65°C in a reactor with immobilized enzymes (Kim *et al.* 2003a; Oh *et al.* 2001), the poor thermal stability of L-AI will limit the reactor's operation period. The reactions at temperatures above 80°C would result in browning as well as isomerization. Although hyperthermophilic L-AIs are more thermostable than mesophilic and moderate thermophilic L-AIs, their application in non-commercial tagatose manufacture might not be accepted in the food industry because isomerases from hyperthermophiles have cobalt as their cofactor (Kim *et al*. 2002, 2003b; Lee *et al*. 2004; Vieille *et al*. 1995, 2001). Additionally, the optimal pH for galactose isomerization is between 7.5 and 8.5. To minimize non-specific side reactions, commercial L-AI needs an acidic pH range, and site-directed mutagenesis and molecular evolution could offer excellent tools to optimize pH effectively (Sriprapundh *et al*. 2003; van Tilbeurgh *et al*. 1992). Also, the solved 3D structures of L-AI may indicate potential mutation targets on or near the active site for altering the substrate feature, the pH activity range, and the increasing understanding of the isomerization reaction mechanism.

To meet the optimal properties of L-AI for commercial production of D-tagatose, molecular breeding was also applied to generate recombined genes by homologues

recombination thorough spontaneous molecular breeding, resulting in progenitor hybrid L-AIs with combined properties from parental sources from L-AI genes encoding hexameric L-AIs of mesophile *Escherichia coli* AI (ECAI), thermophile *Geobacillus kaustophilius* AI (GKAI), thermoacidophile *Alicyclobacillus acidocaldarius* AI (AAAI), and thermoalkalophile *Alicyclobacillus sp*. TP-7 AI (ASAI) (Lee *at al*., 2012; Choi *et al*. 2016; Lee *et al*., 2004; Lee *et al*., 2005).

In previous studies, the crystal structure of L-AI from *Escherichia coli* (ECAI) was described at a molecular level (Manjasetty *et al.*, 2006) (Figure 3). Biophysical and electron microscopy studies predicted native ECAI exists as hexamer form both in crystals and solutions (Patrick & Lee, 1969, Wallace *et al*., 1978. Pauley *et al*., 1972), while thermophilic L-AIs exists as a tetramer in solution (Lee *et al*., 2004). The comparison of structure and sequence alignment of several L-AIs showed the conservation of residues in the subunit interactions mesophilic, thermophilic, and hyperthermophilic enzymes (Manjasetty *et al.*, 2006). The tagatose conversion of mesophilic ECAI is 34% at 35°C (Roh *et al*., 2000). Other studies also revealed the contribution of Mn^{2+} in intrinsic activity and thermostability of ECAI and suggested the modification in hexamer forming related to manganous ion (James W. Patrick, 1971). A kinetic study of another mesophilic L-AI from *Bacillus halodurans* (BHAI) showed that the catalysis does not depend on Mn^{2+} presence (Lee, Choe, *et al.*, 2005a). In contrast, divalent metal ions appear to be an absolute requirement for the catalytic activity of thermophilic L-AIs; furthermore, these types of L-AIs are more stable in the presence of Mn^{2+} (Lee, Choe, *et al.*, 2005a). Indeed, *T. neapolitana* AI (TNAI) and *T. maritima* AI (TMAI) isomerized D-galactose to D-tagatose at 80°C in the presence of divalent metal ions such as Co^{2+} or Mn^{2+} with yields of > 60% (Jorgensen *et al.*, 2004; Izumori *et al.*, 1988). Notably, such (hyper)thermophilic L-AIs were found to be highly dependent on metal ions for both activity and structural stability, which distinguished them from their mesophilic counterparts (Patrick *et al.*, 1971). Residues at the interface between the subunits of multimeric L-AI were essential for the catalytic hole in bioinformatics-led studies.

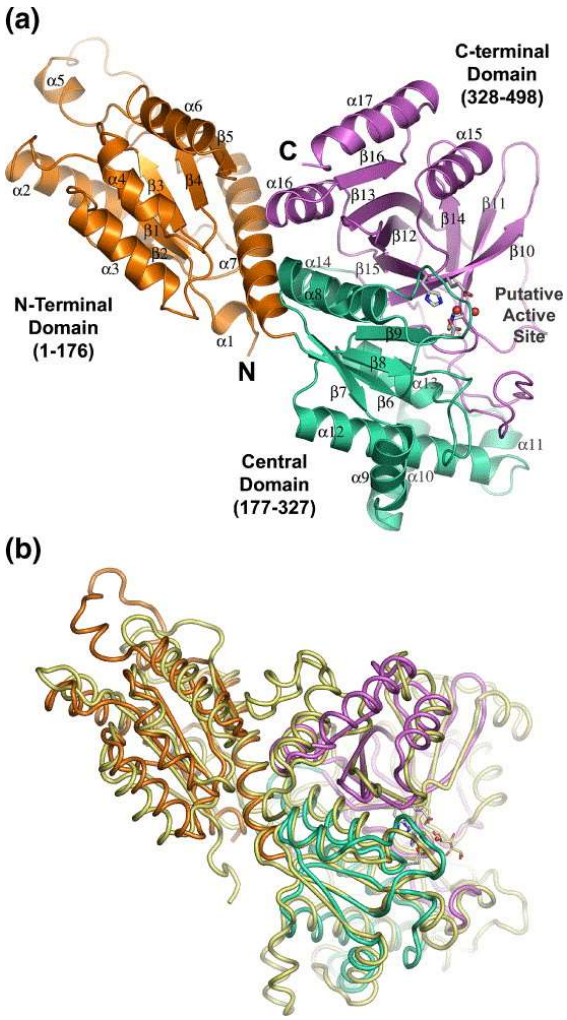


Figure 3: ECAI monomer and its structural homolog ECAI. (Manjasetty *et al.*, 2006)

(a) Ribbon diagram of ECAI monomer colored by domains and secondary structure elements such as domains, N terminus, and C terminus labeled.

(b) Superposition of the Cα trace of monomers: a subunit of ECAI isomerase colored by domains and ECFI (PDB ID, 1FUI) illustrated in yellow.

Secondly, GKAI exist as a hexamer with six identical subunits (Choi *et al*.,2016). In Figure 4, a ribbon representation of the monomer structure of GKAI is exhibited. Although advantages of (hyper)thermophilic L-AIs for the D-tagatose commercial producing from D-galactose are known well, their substrate characteristics and thermostabilities, directly related to conversion performance, need to be more developed through protein engineering for the cost-effectiveness of biological process. In-depth observation of crystal structures gradually discloses the common principles of the characteristic structural L-AI conformation and the molecular mechanism for metal-mediated catalysis and thermostability through interactions of subunits. Both isomerizations can occur in aqueous environments, with Mn^{2+} ion playing a role in stabilizing transition-state for two reactions, implying that metal ion catalysis of isomerization gives a simple interpretation of GKAI's need for divalent metal ions. It was proposed that the catalytic activity of thermostable L-AIs was highly reliant on metal ions. Furthermore, thermophilic L-AI experiences conformational changes induced by Mn^{2+} at enhanced temperatures opposite to mesophilic L-AIs. Indeed, the crystal structures of GKAI in the apo and holo forms with and without bound L-arabitol reveal that L-AIs are metal-dependent for catalysis. Each monomer of hexameric GKAI undergoes significant transpositions of some amino acids caused by metal and substrate binding, supporting previous discoveries that the binding of Mn^{2+} to GKAI transforms the inactive state of the enzyme into an active form at enhanced temperatures. Together, GKAI may exhibit sequential, ordered kinetics, with Mn^{2+} binding first, followed by L-arabinose (Choi *et al*.,2016).

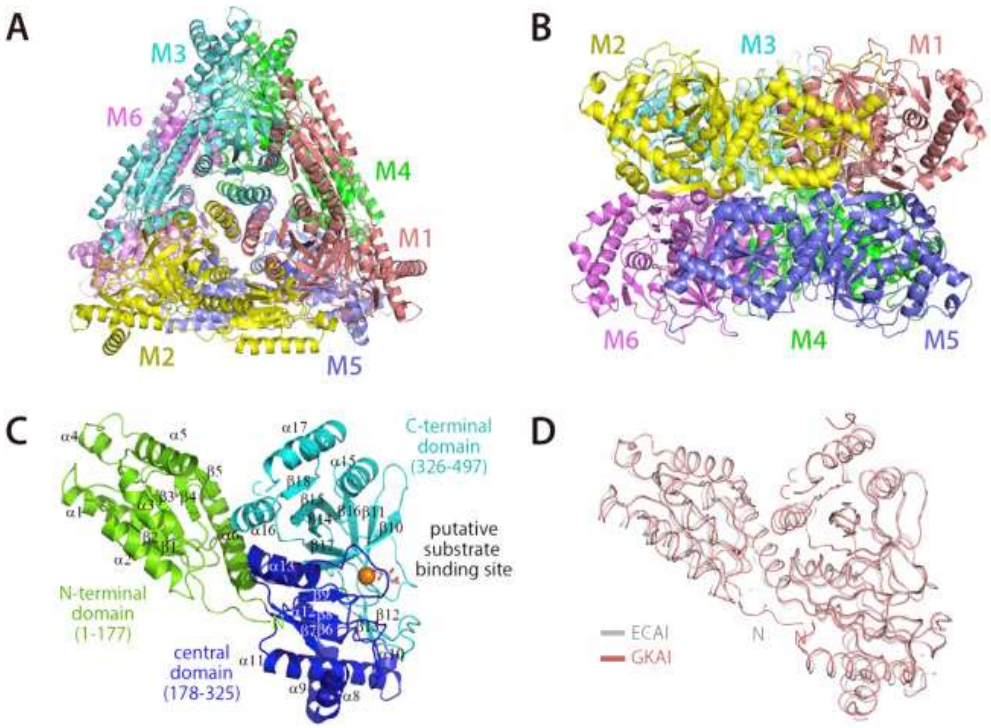


Figure 4: Overall structure of GKAI. (Choi *et al*.,2016)

(A) Top view and (B) side view of the hexamer structure. The six molecules are colored in salmon (M1), yellow (M2), aquamarine (M3), green (M4), slate (M5), and violet (M6).

(C) Ribbon representation of GKAI structure in complex with Mn^{2+}. The structure of GKAI consists of the N-terminal domain shown in green, the central domain in cyan, and the C-terminal domain in blue. The small orange sphere indicates a metal ion.

(D) The apo-GKAI (4R1O) monomer is superimposed on the apo-ECAI monomer (2AJT) colored in deep salmon and white, respectively.

*Alicyclobacillus acidocaldarius (*former name is *Bacillus acidocaldarius*) is a thermoacidophilic bacterium that was isolated by Darland and Brock (1971), from acid thermal environments, both aqueous and terrestrial, at a temperature from 45°C up to 70°C (optimally 60°C to 65°C) and pH values from 2 to 6 (optimally 3 to 4) (Darland & Brock, 1971, Wisotzkey *et al*., 1992). To date, several acidic proteins from *Alicyclobacillus acidocaldarius* have been investigated for pH-related properties. While L-AIs from *Bacillus species*, including *B.subtilis* and *B. halodurans* and *Geobacillus stearothermophilus* have high conservation in amino acid sequences (60%) had different pH optimum ranges. As a result, L-AI from *A. acidocaldarius* may be found as an ideal model to explore critical amino acid residues for acid tolerance of L-AIs. Based on the DNA sequence of L-AI, the 1,494-bp araA gene of *A. acidocaldarius* encoded a polypeptide of 497-amino-acid with an estimated molecular mass of 56 kDa.

Over 200 thermoacidophiles *Alicyclobacillus sp*. strains originated from hot springs in Indonesia were grown on EM-1 medium added with L-arabinose at pH 2 and 65°C, and then 20 strains were isolated and performed L-AI activity assay. TP7 with the highest L-AI activity was chosen from these results, and its 16S rRNA gene sequence was determined. According to the sequencing data, the similarity of TP7 with *Alicyclobacillus* strains was from 97% to 99%. Hence, this isolated strain was identified as *Alicyclobacillus sp.* TP7 (Korean Collection for Type Cultures 33088) and the 1,485-bp araA gene encoded a 494-amino-acid polypeptide (*Alicyclobacillus sp.* TP7 AI [ASAI]). Also, the amino acid sequence of ASAI was found to be highly similar (68% to 96%) to that of *A. acidocaldarius* and *Geobacillus stearothermophilus*.

With the above background of L-AIs, from the scientific and industrial point of view, the governing principles for distinctive physico-chemical characteristic properties of each L-AI are still unanswered. Even further, the relationship between those physical properties such as pH, temperature and catalytic efficiencies, and structural characteristics has not been investigated. Knowing the facts is not only crucial to expanding the horizon of L-AI research but also to apply the knowledge to industrial mass production of tagatose.

Significantly, the latter is likely to be possible by an engineered L-AIs, which is designed to produce tagatose in larger amounts under the industrially beneficial production conditions, including optimized pH, temperature, and substrate specificity. To solve this problem, it is necessary to investigate the domain structure-function relationship of each distinctive L-AI. Therefore, I focused on the structure determination of the hybrid L-AIs that showed enhanced galactose conversion activities under combinatorial physico-chemical conditions. In the line of accomplishing the goal, I attempted to overexpress, purify and crystallize the hybrid L-AIs. Furthermore, the preliminary X-ray diffraction analysis of hybrid L-AIs has also been performed.

II. MATERIALS AND METHODS

1. Plasmid constructs

The plasmid constructs containing the genes of interest hybrid AI-8 and AI-10 were cloned into pET-28a (+) provided by Dr. Dong-Woo Lee (Department of Biotechnology, Yonsei University, Korea). For overexpression, the plasmids were transferred into *E.coli* BL21 (DE3).

The structure of hybrid AI-8 comprises the 4 domains of thermophilic GKAI (aa 1-125), thermoalkalophilic ASAI (aa 126-284), thermoacidophilic AAAI (aa 285-378), and thermophilic GKAI (aa 379-497), respectively. The structure of hybrid AI-10 consists of the first domain from thermoalkalophilic ASAI (aa 1-123), the second from thermophilic GKAI (aa 124-282), the third from mesophilic ECAI (aa 283-379), and the fourth from thermophilic GKAI (aa 380-498) (Figure 14).

2. Over-expression, Purification, and Isolation

The transformed cells were grown in Luria-Bertani (LB) medium with kanamycin as a selection marker. The recombinant protein was overexpressed by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at the mid-exponential phase (A_{600} = 0.6) and continued to grow at 37°C for 5 hours. The cells were pelleted and frozen at -80°C. For purification, the frozen bacterial pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, and 10 mM Imidazole) supplemented with 1 mM of phenylmethylsulfonyl fluoride (PMSF) as a protease inhibitor, 1 mM of dithiothreitol (DTT), DNase, and Lysozyme. Cell disruption was carried out by sonication on ice (Amplitude 35% for 7 minutes with an alternative pulse on/off 20 seconds), and the lysates were centrifuged at 22,000 g for 30 minutes at 4°C to separate insoluble cell debris. The supernatant was used as crude cell extract and further purified using Ni-NTA agarose columns.

Ni-NTA agarose (from Qiagen GmbH, Germany) is a nickel-nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography matrix for recombinant protein purification with 6 consecutive tagged histidine residues (6xHis-tagged). The columns were packed with 3 mL of Ni-NTA agarose suspension (in 20% ethanol) and equilibrated with 2 column volumes of lysis buffer. The supernatants were loaded onto these columns. Using a peristaltic pump (Bio-Rad, USA), the speed of the fluid stream was set to 1 mL min^{-1}. The 5 column volumes of washing buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl and 20 mM Imidazole) were used to wash the column, and the speed of the fluid stream was set to 2 mL min^{-1}. The target protein was eluted and collected with each of 1 mL in the elution buffer (20 mM Tris-HCl pH 7.5, 500 mM NaCl, and 250 mM Imidazole). For each step, 40 µl samples were saved and run on 10% Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Gels were visualized by Coomassie Brilliant Blue R250 staining (Wako, India).

The eluted solution was dialyzed twice overnight against 2 L of dialysis buffer (20 mM Tris-HCl pH 7.5 and 150 mM NaCl) at room temperature to reduce the high concentration of salt and imidazole from the protein solution. 2 mM calcium chloride and 1 mM manganese chloride were supplemented to the protein solution followed by the addition of human-thrombin (Haematologic Technologies Inc., USA) with 35 units for 10 mg of fusion proteins to cleave the His-tag out of the target protein.

3. Size exclusion chromatography

The target proteins were further purified using size-exclusion chromatography on a HiLoad 16/60 Superdex 200 pg column (GE Healthcare, USA) equilibrated with crystallization buffer of 20 mM Tris-HCl and 50 mM NaCl, pH 7.4. The fractions from the elution peak of size exclusion chromatography were collected, and the purity of the protein was examined by SDS-PAGE and envisioned by Coomassie Brilliant Blue R250. The column was calibrated using Ferritin (440 kDa), Aldolase (158 kDa), Conalbumin (75 kDa), Ovalbumin (44 kDa), Carbonic Anhydrase (29 kDa), and Ribonuclease A (13.7

kDa). The standard curve was obtained by a plot of K_{av} against molecular weight (M_r), which was used to calculate the oligomeric state of the L-AIs. The protein concentration was measured by UV-Vis spectrophotometer **(**Thermo Scientific™ NanoDrop 2000 and 2000c).

4. Crystallization

The purified AI-8 and AI-10 were concentrated to 10 µg/µl in crystallization buffer (20 mM Tris HCl pH 7.5, 50 mM NaCl) using a 10 kDa cut-off concentrator (Amicon® Ultra Centrifugal Filter, Merck, Germany). Hanging-drop vapor diffusion crystallization trials (1 µl of protein solution and 1 µl of reservoir solution) were on a VDX plate of 24 wells (Hampton Research, USA) by using Hampton Research crystallization kits (such as Crystal Screen, Crystal Screen 2, Crystal Screen Lite, PEG/Ion, PEG/Ion 2, Index and SaltRx, USA), Rigaku kits (Wizard Classic I, II, III and IV, Japan), and Molecular Dimensions (Midasplus 1, 2, USA). All plates were placed in a 20°C incubator and 4°C cold room, and the time point that the crystals emerged were recorded carefully. For the X-ray diffraction experiments, crystals in the various shapes and angles were chosen and flash-cooled immediately in the presence of pre-screened cryoprotectant, including 25% of glycerol and 10% of DMSO to the reservoir solution.

5. X-ray crystallography

The crystals of AI-8 and AI-10 were subjected to X-ray diffraction in Pohang Light Source (PLS) beamline 7A (Pohang, South Korea) using ADSC Q270 detector. To obtain the optimal data collection parameters, a couple of exposure tests were done at two or a set of orthogonal orientations, such as 0°, 90°, 180°, 270°, and 360°. Subsequently, X-ray diffraction was performed at -173°C in a cryogenic N_2 gas stream by the rotation method with an oscillation of 1.0° for 1-sec exposure over a 360° range at a wavelength of 0.97942 Å. Using the HKL-2000 program (HKL Research Inc., USA), these diffraction datasets were processed in four distinct major stages: spot searching followed by autoindexing, parameter refinement, integration, and scaling.

6. Phasing, Structure Determination, and Refinement

The initial structure of two hybrids was determined by molecular replacement using MOLREP program in CCP4 (Emsley & Cowtan, 2004). The structure of apo-GKAI (PDB code 4R1P) was a reference model for hybrid AI-8 with a percentage identity of 87.93%. The structure of apo-ECAI (PDB code 2AJT) was used as a search model for hybrid AI-10 with 65.31% identity. Subsequent model building process and refinement using Refmac5 (Murshudov *et al*., 2011) reduced the R_{factor}/R_{free} values to the reliable range for the protein structural model.

III. RESULTS AND DISCUSSION

1. Primary sequence analysis of hybrid L-arabinose isomerases, AI-8 and AI-10

The selected hybrid L-AIs (AI-8 and AI-10) were sequenced and compared with parental L-AIs (Introduction and Materials and Methods section) (Figure 5). Four well-characterized and distinct homologs exhibit a wide range of the temperature and pH dependence of catalysis and stability from 40°C−90°C and 4−10, respectively, as well as specific promiscuous activity toward C5 and C6 sugars in their specific clans. Among them, two hybrid constructs with different temperature and pH dependence from their parental enzymes, including AI-8 and AI-10, were chosen for this structure study. Subsequent structure analysis of these hybrids can reveal the proposed adaptive traits that contribute to the temperature and pH dependence of L-AIs. The structures of ECAI (Manjasetty *et al*., 2006) and GKAI (from our lab, Choi *et al*., 2016) were known and used as search models for the hybrid AI-8 and AI-10 in this report.

Alignment of the amino acid sequences of L-AIs (*Alicyclobacillus acidocaldarius* AI (AAAI), *Alicyclobacillus sp.* TP7 AI (ASAI), *Escherichia coli* AI (ECAI), *Geobacillus kaustophilius* AI (GKAI), *Thermotoga maritima* AI (TMAI), and two hybrid AI-8 and AI-10) exhibited high levels of sequence identity (similarity) of >54% (68%) (Figure 5).
The 10 conserved regions containing at least 5 consecutive identical amino acid residues in the sequences were found. The residues of active sites on model structures of GKAI (4R1P) and ECAI (2AJT) were found in these high conserved sequence regions. Although the sequences of hybrid AI-8 and AI-10 are not much different from their parental sources, they possess outstanding characteristics such as a wide range of temperature and pH for catalysis and stability. Therefore, the structure research of hybrid L-AIs can confirm the identical active sites and imply the adaptive evolution of homologous enzymes fitting into harsh environments.

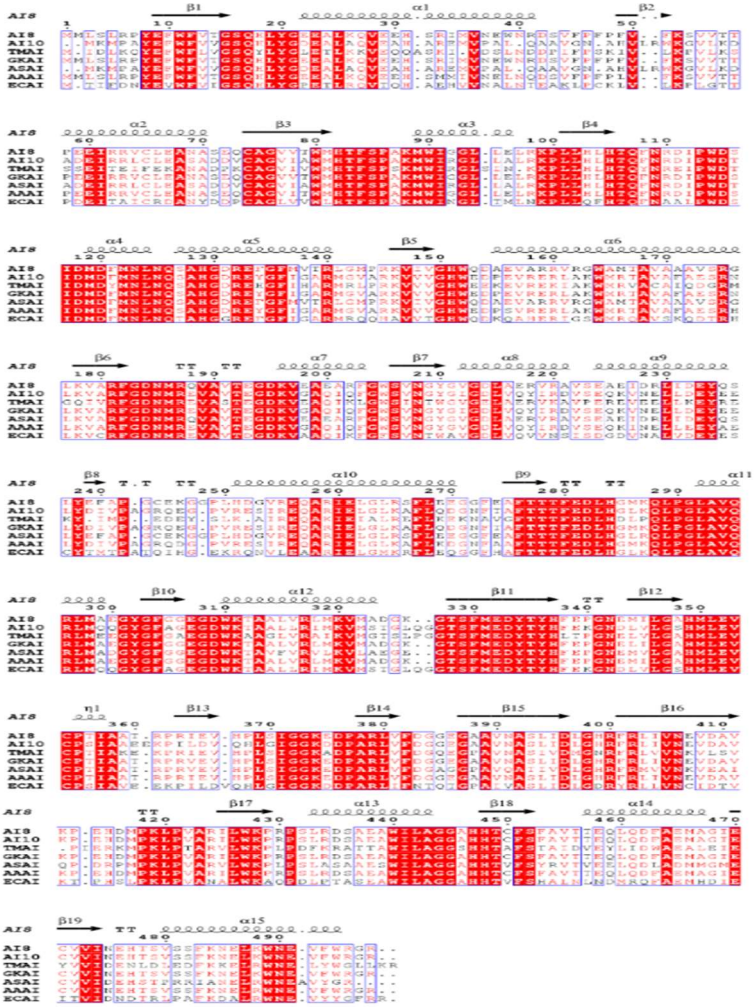


Figure 5: **Multiple amino acid sequence alignment of L-AI proteins.** Sequence alignments of *Alicyclobacillus acidocaldarius* AI (AAAI), *Alicyclobacillus sp.* TP7 AI (ASAI), *Escherichia coli* AI (ECAI), *Geobacillus kaustophilius* AI (GKAI), *Thermotoga maritima* AI (TMAI) and two hybrid AI-8 and AI-10. The alignment was performed using *ESPript 3*. Conserved residues are surrounded in navy boxes separately. Totally conserved residues among the seven molecules are highlighted with a red background, and less conserved residues are shown in red letters.

2. Expression and Purification

The recombinant protein expression test in the small scale was firstly carried out to determine the optimal concentration of IPTG, time, and temperature for maximal protein production. The IPTG concentrations were tested between 0.5 mM and 1 mM, and the temperature for the induction step were tested at 37°C and 18°C overnight. The results proved that the more soluble protein yield was obtained at 37°C for 5 hours, and 1 mM IPTG was added when an OD_{600} reached ~ 0.6 (Lane 1-4 of Figure 6a-AI-8 and 7a-AI-10). The Ni^{2+} chelate affinity chromatography was used to purify the proteins in soluble form with a regular molecular weight (M_r) of ~ 55 kDa (Lane 7 of Figure 6a- AI-8 and 7a-AI-10). The thrombin treatment was followed to remove the extra 6x His-tag, and the M_r of the intact proteins were estimated to be ~ 53 kDa by SDS-PAGE (Figure 6b-AI-8 and 7b-AI-10). However, the size exclusion chromatography (SEC) analysis indicated that the apo form of hybrid AI-8 and AI-10, as in other L-AIs, existed as a hexamer with M_r of ~ 287.037 kDa and ~ 289.801 kDa, respectively (Figure 8a-AI-8 and 9a-AI-10). The SEC fractions were checked on SDS gel (Figure 8b-AI-8 and 9b-AI-10) and pooled together for the crystallization trial.

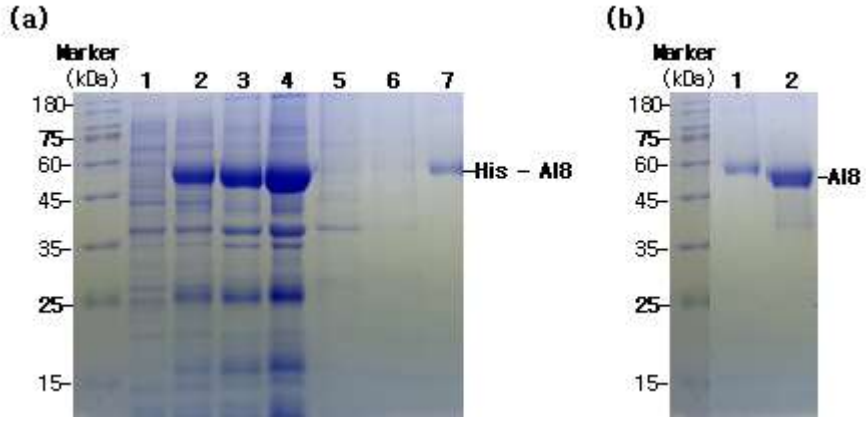


Figure 6: SDS-PAGE analysis of purified 6xHis-tagged AI-8 (a) and intact AI-8 (b)

(a) The analysis on 10% SDS-PAGE demonstrates the step-by-step purification process. The first lane is the molecular weight ladder (kDa). Lane 1 is the whole cell before induction. Lane 2 is the whole-cell after induction. Lane 3 is the pellet from the cell extract. Lane 4 is the supernatant from the cell extract. Lane 5 is the Ni-affinity unbound fraction. Lane 6 is the Ni-NTA agarose column wash. Lane 7 is purified 6xHis-tagged AI-8.

(b) The extra 17 amino acid residues containing 6x His-tagged were cleaved by thrombin enzyme. The first lane is the molecular weight ladder (kDa). Lane 1 is 6xHis-tagged AI-8 (~55 kDa). Lane 2 is Thrombin digested AI-8 (~53 kDa).

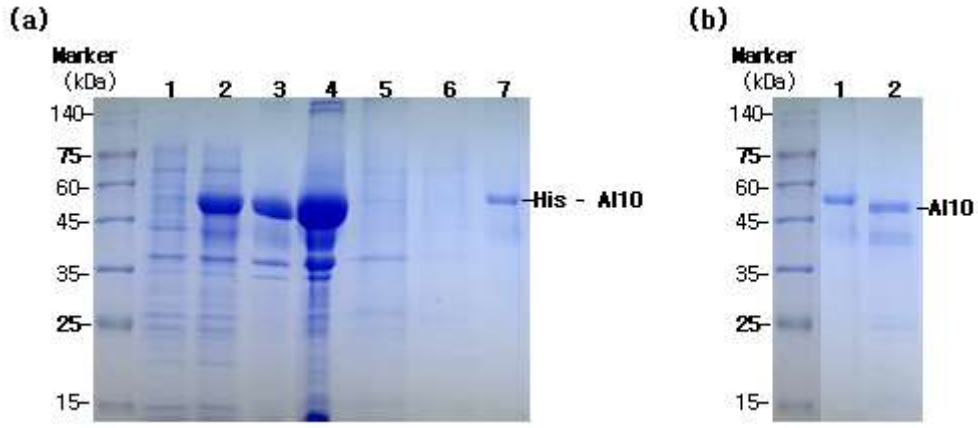


Figure 7: SDS-PAGE analysis of purified 6xHis-tagged AI-10 (a) and intact AI-10 (b)

(a) The analysis on 10% SDS-PAGE demonstrates the step-by-step purification process. The first lane is the molecular weight ladder (kDa). Lane 1 is the whole cell before induction. Lane 2 is the whole-cell after induction. Lane 3 is the pellet from the cell extract. Lane 4 is the supernatant from the cell extract. Lane 5 is the Ni-affinity unbound fraction. Lane 6 is the Ni-NTA agarose column wash. Lane 7 is purified 6xHis-tagged AI-10.

(b) The extra 17 amino acid residues containing 6xHis-tagged were cleaved by thrombin enzyme. The first lane is the molecular weight ladder (kDa). Lane 1 is 6xHis-tagged AI-10 (~55 kDa). Lane 2 is Thrombin digested AI-10 (~53 kDa).

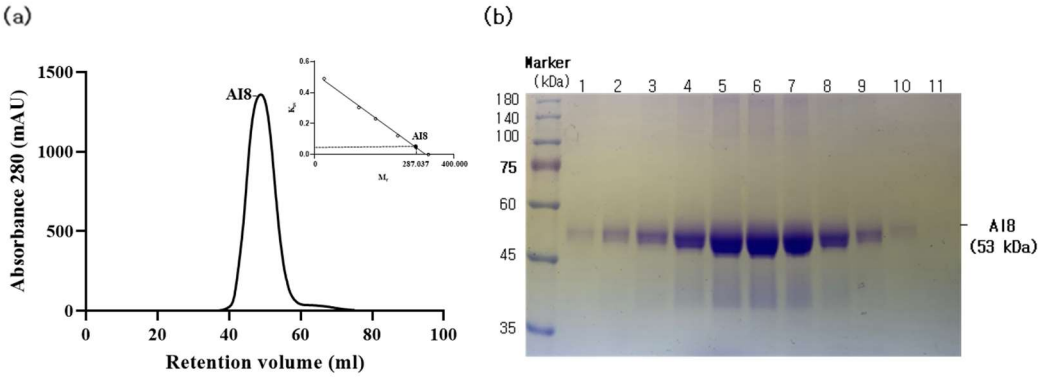


Figure 8: Size-exclusion chromatography of hybrid AI-8

(a) The analytical size-exclusion chromatography (SEC) of AI-8 is performed on the ÄKTA FPLC Purifier System with a packed HiLoadTM 16/600 SuperdexTM 200 prep grade column. The protein sample was detected in one peak, and all fractions of this peak were then collected and examined on 10% SDS-PAGE. The apo form of hybrid AI-8 exists as a hexameric (M_r ~287.037 kDa).

(b) The purity of AI-8 after SEC was analyzed by 10% SDS-PAGE. Lane from number 1 to number 11 stand for fractions from B6 to C1, respectively. All high-purity fractions from number 1 to number 11 were then pooled and used to set up crystallization.

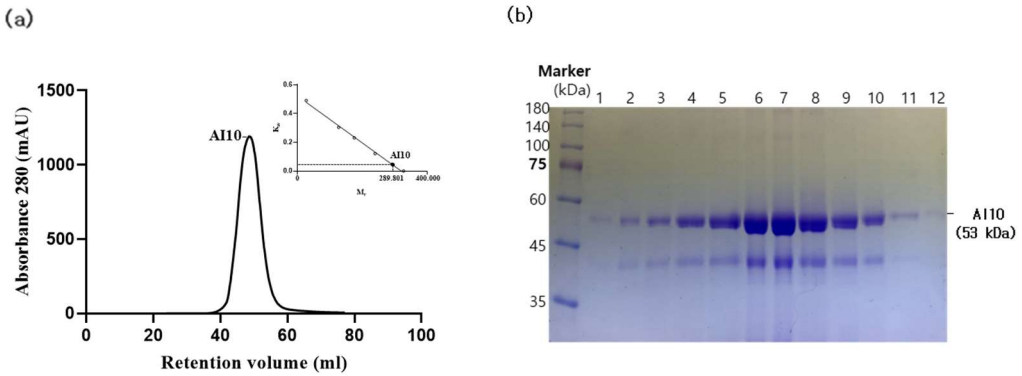


Figure 9: Size-exclusion chromatography of hybrid AI-10

(a) The analytical size-exclusion chromatography (SEC) of AI-10 is performed on the ÄKTA FPLC Purifier System with a packed HiLoadTM 16/600 SuperdexTM 200 prep grade column. The protein sample was detected in one peak, and all fractions of this peak were then collected and examined on 10% SDS-PAGE. The apo form of hybrid AI-10 exists as a hexameric (M_r ~289.801 kDa).

(b) The purity of AI-10 after SEC was analyzed by 10% SDS-PAGE. Lane from number 1 to number 12 stand for fractions from B6 to C1, respectively. All high-purity fractions from number 1 to number 12 were then pooled and used to set up crystallization.

3. Crystallization and Data collection

The chimeric L-AIs were concentrated to 10 µg/µl by a 10 kDa centrifugal concentrator (Amicon® Ultra Centrifugal Filter, Merck, Germany) in crystallization buffer with 1 mM manganese (II) chloride. The protein was crystallized by the hanging-drop vapor diffusion method using various crystallization kits (Hampton, USA, and Wizard, USA). Crystals of the hybrid AI-8 and AI-10 were grown from several crystallization conditions (Figure 10 and 11) after 4-7 days at 20°C and 2-4 weeks at 4°C, respectively, and diffracted to the low resolution ranging from 6 to 8 Å, implying that crystal quality needs to be necessarily improved.

For X-ray diffraction experiments, the selective crystals were picked up and arrested in the nylon cryo-loops, soaked in cryo solution followed by flash freezing them using liquid nitrogen at a cryogenic temperature of - 196°C. The apo form of AI-8 crystal, which was obtained from 35% (v/v) (+/-)-2-Methyl-2,4-pentanediol, 100mM Imidazole/ Hydrochloric acid pH 8.0, 25% glycerol at 4°C (Table 1), diffracted to 2.5 Å (Figure 12) on beamline 7A at Pohang Light Source (PLS) equipped with ADSC Q270 detector. The diffraction data set was then indexed, integrated, and scaled by using HKL2000. The apo form AI-8 belonged to space group C2, with unit-cell dimensions a = 204.59, b = 81.91, c = 192 Å, $\alpha = \gamma = 90$, $\beta = 117.9°$. The calculated Matthews coefficient of apo AI-8 form (Matthews, 1968; Kantardjieff & Rupp, 2003) was 2.24 $Å^3\ Da^{-1}$ with a solvent content of 45%, indicating that the asymmetric unit of the crystal is likely to contain six molecules.

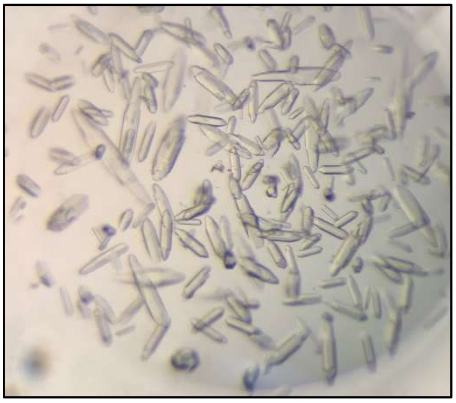
The adonitol-bound AI-8 crystal was also generated by soaking the AI-8 crystal with 10 mM adonitol in crystallization solution for 5 seconds and diffracted to 2.6 Å. The holo form (Adonitol-bound) of hybrid AI-8 belonged to space group C2 with the unit cell parameters a = 205.58, b = 82.12, c = 192.73 Å, $\alpha = \gamma = 90$, $\beta = 117.93°$. The Matthews coefficient parameter of 2.26 $Å^3\ Da^{-1}$ indicated the presence of six molecules per asymmetric unit of adonitol-bound AI-8 crystal, which corresponded to 46% solvent

content (Table 1 and Figure 12).

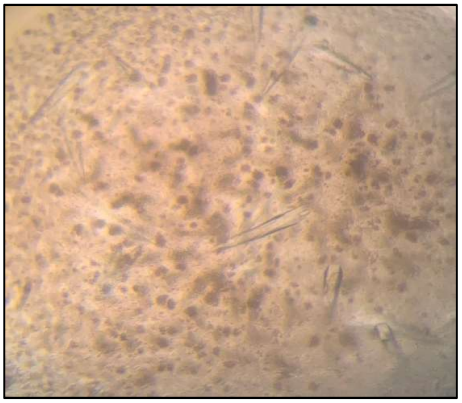
The apo form of AI-10 crystal was obtained from the crystallization solution containing 30% (v/v) Polyethylene glycol 400, 100mM Sodium Cacodylate/ Hydrochloric acid pH 6.5, and 200 mM Lithium sulfate at 4°C. The crystal diffracted to 3.2 Å resolution when 25% glycerol and 10% DMSO mixed with mother liquor solution was used as a cryo-protectant. The apo form of hybrid AI-10 belonged to space group $C222_1$ with unit cell dimension a = 148.48, b = 258.58, c = 165.73 Å, α = β = γ = 90°. The Matthews coefficient parameter of 2.5 $Å^3$ Da^{-1} indicated that the asymmetric unit of AI-10 crystal contained six molecules, which corresponded to a calculated solvent content of 51% (Table 1 and Figure 12).

All three types of best diffracting crystals were obtained at 4°C, which might imply that the relatively slow nucleation and growth of the crystals are essential requirements for the proper crystal packing processes for L-AIs.

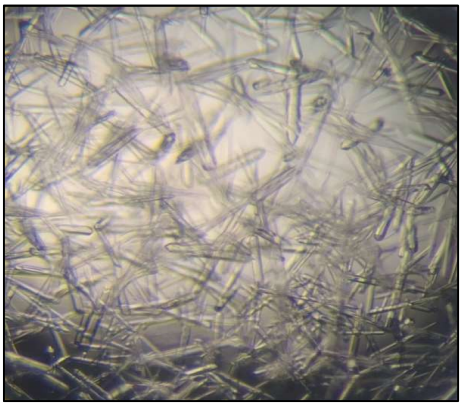
Figure 10: Crystals of hybrid AI-8 obtained from initial screening conditions



0.2 M Potassium sodium tartrate tetrahydrate
0.1 M BIS-TRIS pH 6.5
10% w/v Polyethylene glycol 10,000



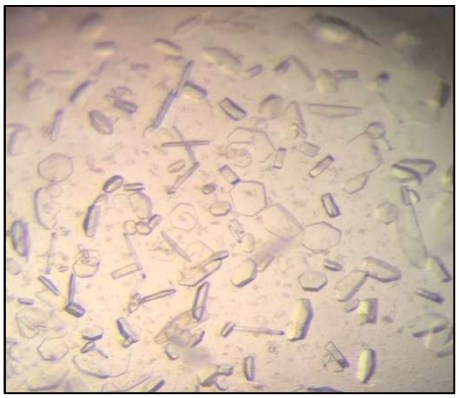
0.1 M MES monohydrate pH 6.5
12% w/v Polyethylene glycol 20,000



0.1 M Sodium acetate trihydrate pH 7.0
12% w/v Polyethylene glycol 3,350

0.1 M HEPES 7
0.2 M Ammonium acetate
25% w/v SOKALAN® HP 56

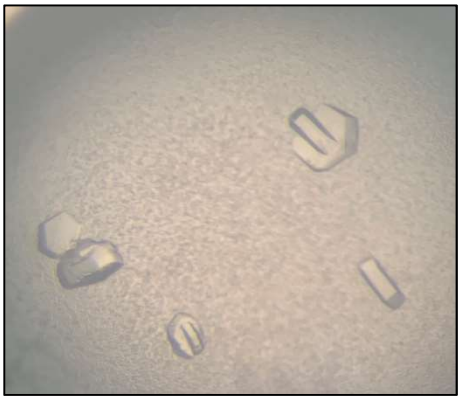
Figure 11: Crystals of hybrid AI-10 obtained from initial screening conditions



0.1 M BICINE pH 9.0
2% v/v 1,4-Dioxane
10% w/v Polyethylene glycol 20,000

0.1 M Bis-Tris pH 6
15 % w/v SOKALAN® CP 5

1% w/v Tryptone
0.001 M Sodium azide
0.05 M HEPES sodium pH 7.0
12% w/v Polyethylene glycol 3,350

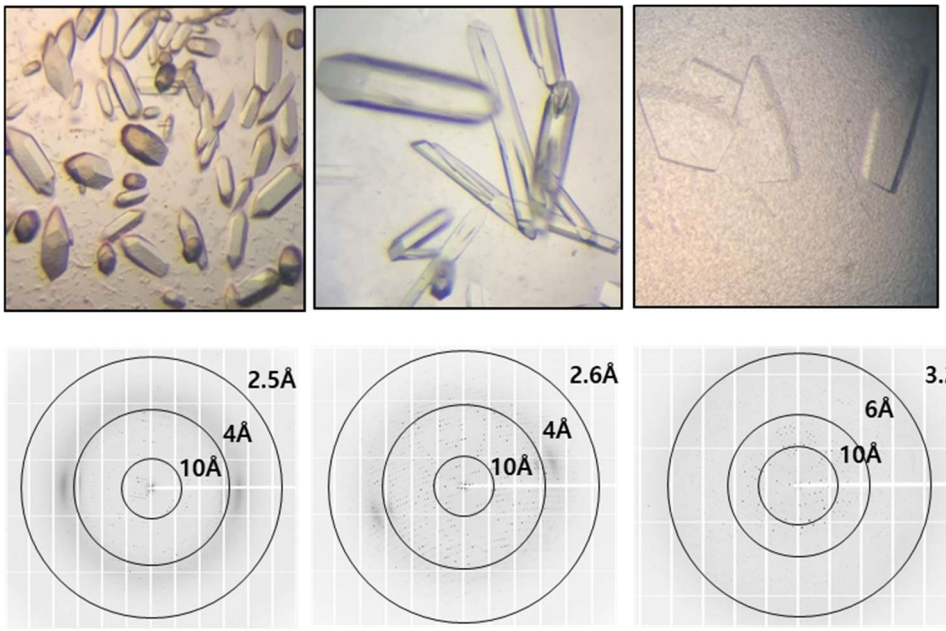


0.1 M Imidazole pH 7.0
12% w/v Polyethylene glycol 20,000

Table 1: Crystallization conditions of best hybrid AI-8 and AI-10 crystals

Crystal	Buffer	Precipitant 1	Precipitant 2	Additive	Temperature
Hybrid AI-8 (Apo)	100mM Imidazole/ Hydrochloric acid pH 8.0		35% (v/v) (+/-)-2-Methyl-2,4-pentanediol	25% Glycerol	4ºC
Hybrid AI-8 (With Adonitol)	100mM Imidazole/ Hydrochloric acid pH 8.0		35% (v/v) (+/-)-2-Methyl-2,4-pentanediol	10mM Adonitol	4ºC
Hybrid AI-10 (Apo)	100mM Sodium cacodylate/ Hydrochloric acid pH 6.5	200mM Lithium sulfate	30% (v/v) Polyethylene glycol 400	25% Glycerol 10% DMSO	4ºC

Figure 12: Best quality crystals with high resolutions of hybrid AI-8 and AI-10



Hybrid AI-8 (Apo)	Hybrid AI-8 (with Adonitol)	Hybrid AI-10 (Apo)

Table 2: Crystallographic data collection statistics for hybrid AI-8 and AI-10

Diffraction statistics	Hybrid AI-8 (Apo)	Hybrid AI-8 (With Adonitol)	Hybrid AI-10 (Apo)
Beamline	PAL7A	PAL7A	PAL7A
Wavelength (Å)	0.97942	0.97942	0.97942
Temperature (K)	100	100	100
Resolution(Å)	50-2.50 (2.54-2.50)	50-2.60 (2.64-2.60)	50-3.20 (3.26-3.20)
Space group	C2	C2	$C222_1$
Unit cell a, b, c (Å)	α=204.59 b=81.91 c=192.00	α=205.58 b=82.12 c=192.73	α=148.48 b=258.58 c=165.73
α, β, γ (°)	α=γ=90 β=117.90	α=γ=90 β=117.93	α=β=γ=90
Total reflections	654992	585492	517675
Unique reflections	97080 (4776)	88026 (4409)	54358 (2690)
R_{merge} (%)	12.1 (85.1)	19.5 (99.7)	18.7 (97.5)
Completeness (%)	99.2 (99.5)	99 (98.7)	99.8 (99.9)
Redundancy	6.7 (6.6)	6.7 (6.1)	9.5 (9.2)
Average I/σ(I)	9.5 (1.2)	16.9 (1.8)	4.3 (0.7)
Matthew's coefficient ($Å^3Da^{-1}$)	2.24	2.26	2.5
Solvent content (%)	45	46	51
No. of molecules in the asymmetric unit	6	6	6

Values in parentheses correspond to the highest resolution shell.
R-merge=$\Sigma(I-\langle I\rangle)/\Sigma I$, I and $\langle I\rangle$ are the intensity and the mean value of all the measurements of an individual reflection.

4. Initial structure determination by Molecular replacement

The molecular replacement (MR) using the program MOLREP in the CCP4 suite for the apo form of hybrid AI-8 was obtained using a model of GKAI-4R1P (87.93% identity) with an initial R-factor (R_{free}) of 0.4251 (0.4377), which finally converged to 0.2656 (0.3352) upon manual model building and refinement. The adonitol-bound AI-8 structure resulted in the initial R-factor (R_{free}) of 0.3815 (0.3895) using a search model of GKAI-4R1P (87.93% identity). After the several first rounds of refinement, the R-factor (R_{free}) of adonitol-bound AI-8 structure reached 0.2393 (0.3203). The apo form of AI-10 structure was built using ECAI-2AJT (65.31% identity) as a search model with the initial R-factor (R_{free}) of 0.4445 (0.4509) and the final R-factor (R_{free}) of 0.3116 (0.3953) after some refinement cycles and subsequent model building. Refinement and rebuilding of these three structures are currently ongoing in our lab.

As expected, during model building for the AI-8 soaked in adonitol structure, an additional F_0-F_C maps (in green) was revealed. This map was located very close to the active site containing metal-binding sites. Hence, we predicted that the map might correspond to the substrate analogue adonitol (Figure 13). In addition, another excess map was identified near the conserved residues for metal binding sites as seen on the apo form of AI-8 (Glu307, His448, His349, and Glu332) also in the apo form of AI-10 (Glu305, His449, His349, and Glu332) (Figure 13a and 13b). Hence, based on similar metal-binding sites located in ECAI, GKAI, and other L-AIs, we considered the map to be a metal. When Mn was placed in this position, there were no rejections upon refinement. Similarly, modeling adonitol in the larger electron density map in adonitol-bound AI-8 crystal structure fits perfectly with the least rejections (Figure 13c). As in other L-AIs, the metal was located proximal to the substrate-binding site where adonitol was located. The metal in AI-8 was coordinated by residues (including Glu307, His448, His349, and Glu332), and the substrate was coordinated by metal, as in other L-AIs.

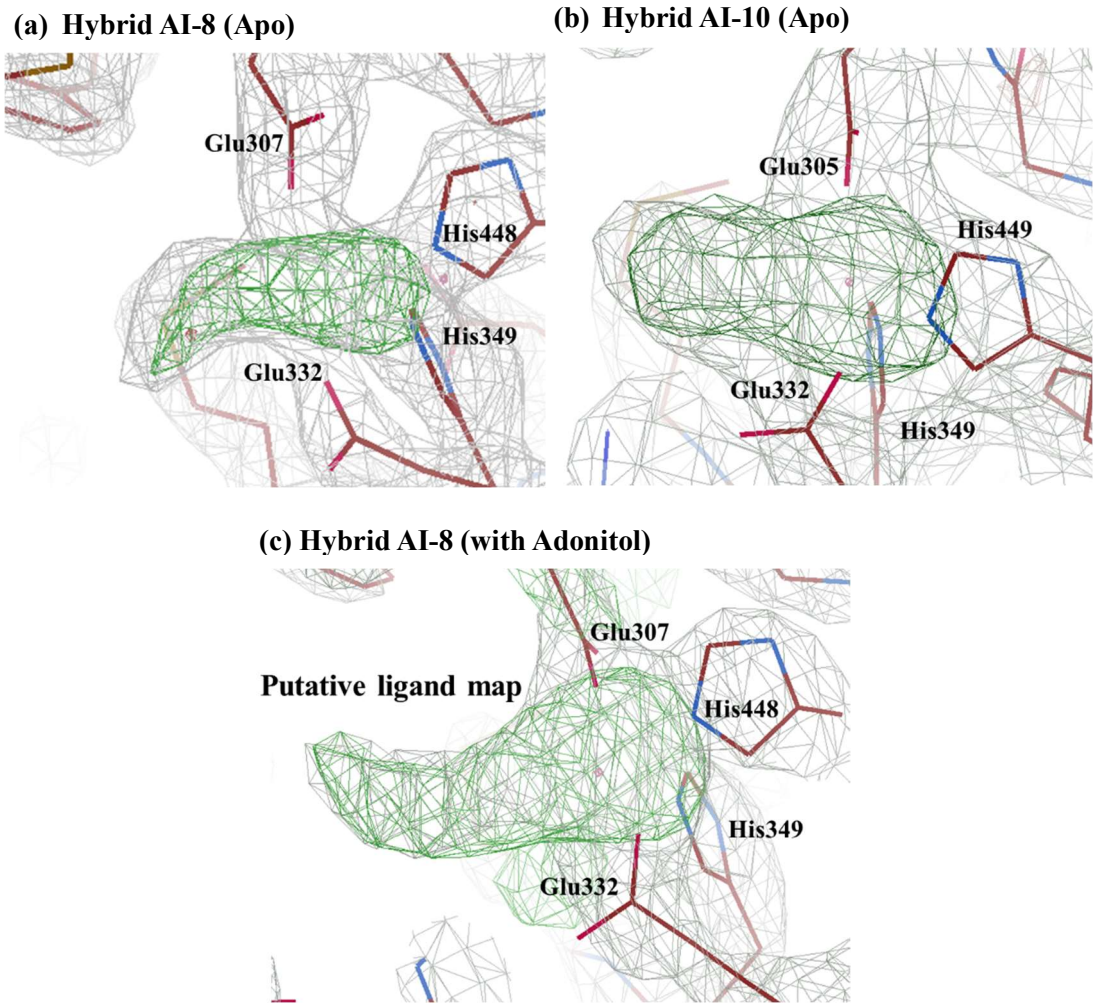


Figure 13: Phasing snapshot for structure determination

The representative snapshot of structures was obtained after the molecular replacement and several refinement cycles. The $2F_0$-F_C maps were illustrated with smooth gray meshes, and the F_0-F_C maps (in green) indicated the presence of ligands in the active sites.

5. Initial structure analysis

The initial structures of chimeric apo-hybrid AI-8, hybrid-AI-8-adonitol, and hybrid AI-10 were determined at resolutions of 2.6 Å, 2.5 Å, and 3.2 Å, respectively and visualized with *PyMOL* V1.8 (DeLano Scientific LLC.) (Figure 14). From the top view, all three solved structures consisted of six subunits (S1, S2, S3, S4, S5, and S6), which formed a sandwich-like hexamer in an asymmetric unit (Figure 14 a and c). The SEC results, together with structural data, strongly suggested that hybrid AI-8 and AI-10 structures may comprise six identical subunits existing as a hexamer.

A ribbon representation of the monomer structure of hybrid AI-8 was shown in Figure 14b, which composed of the first fragment (1-125) of thermophilic GKAI (in blue), the second fragment (126-284) of thermoalkalophilic ASAI (in green), the third fragment (285-378) of thermoacidophilic AAAI (in yellow), and the fourth fragment (379-497) of thermophilic GKAI (in blue). The monomer structure of hybrid AI-10 (Figure 14d) consists of the first fragment (1-123) from thermoalkalophilic ASAI (in green), the second fragment (124-282) from thermophilic GKAI (in blue), the third fragment (283-379) from mesophilic ECAI (in pink), and the fourth fragment (380-498) from thermophilic GKAI (in blue). The in-depth knowledge of how the hybrid structure will enhance the catalytic efficiency compared with the wild-type would be revealed upon complete modeling and refinement.

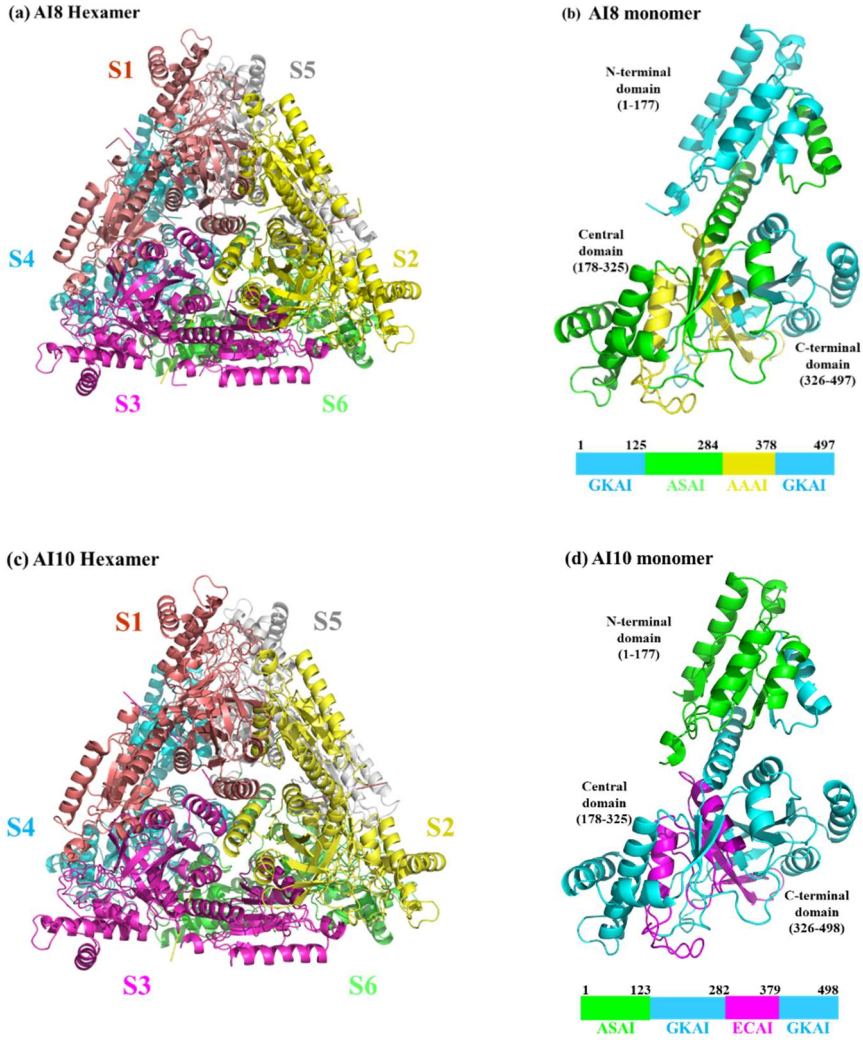


Figure 14: Overall structures of hybrid AI-8 and AI-10

(a) Top view of the hexamer AI-8 structure and **(c)** top view of the hexamer AI-10 structure. The six chains are colored in salmon (S1), yellow (S2), magenta (S3), aquamarine (S4), white (S5), and green (S6). **(b)** The monomer structure of AI-8 includes the segments of GKAI (blue), ASAI (green), and AAAI (yellow), respectively. **(d)** The monomer structure of AI-10 includes the segments of ASAI (green), GKAI (blue), and ECAI (pink), respectively.

IV. CONCLUSION

In this study, the hybrid AI-8 and AI-10 were overexpressed in *E.coli* BL21 (DE3), purified, and crystallized with resolutions of 2.5 Å (AI-8-Apo), 2.6 Å (AI-8-Adonitol), and 3.2 Å (AI-10-Apo). As expected from the diffraction statistics and size-exclusion chromatography, the functional unit of the hybrid L-AIs may be the same as authentic L-AIs, whose oligomeric state is a hexamer. Also, the initial structure reveals that the monomer of each hybrid protein contained the combined fragments of four well-characterized and distinct homologs, resulting in progenitors with combined properties from parental sources. In the active sites from both hybrid L-AIs, the F_0-F_C electron density maps are found, which were coincide with progenitor L-AIs, indicating that the map may correspond to Mn^{2+}. Furthermore, the map of the adonitol, analog of arabinose, is shown from the adonitol-soaking crystals. With those results, it concludes that the hybrid L-AIs obtained by molecular breeding among the distinctive L-AIs may preserve the structural integrity of L-AIs, which is likely to have advantages for industrial application. Currently, further refinement is underway to finalize the complete structures, from which structural domain-specific functions contributing to the temperature and pH dependencies among the L-AIs can be elucidated.

V. REFERENCES

1. Izumori, K. & Tsuzaki, K. (1988). Production of Dtagatose from D-galactitol by *Mycobacterium smegmatis.* J. Ferment. Technol. 66, 225–227.

2. Cheetham PSJ, Wootton AN (1993). Bioconversion of D-galactose into D-tagatose. Enzyme Microb Technol 15:105–108.

3. Manjasetty BA, Chance MR. (2006). Crystal structure of *Escherichia coli* L-arabinose isomerase (ECAI), the putative target of biological tagatose production. J. Mol. Biol. 360:297–309.

4. Roh, H. J., Kim, P., Park, Y. C. & Choi, J. H. (2000). Bioconversion of D-galactose into D-tagatose by expression of L-arabinose isomerase. Biotechnol. Appl. Biochem. 31, 1–4.

5. Patrick, J. W. & Lee, N. (1968). Purification and properties of an L-arabinose isomerase from *Escherichia coli*. J. Biol. Chem. 243, 4312–4318.

6. Levin, G. V. (2002). Tagatose, the new GRAS sweetener, and health product. J. Med. Food, 5, 23–36.

7. Kim, P. (2004). Current studies on biological tagatose production using L-arabinose isomerase: A review and future perspective. Appl. Microbiol. Biotechnol. 65,243–249.

8. Jorgensen, F., Hansen, O. C. & Stougaard, P. (2004). Enzymatic conversion of D-galactose to D-tagatose: Heterologous expression and characterization of a thermostable L-arabinose isomerase from *Thermoanaerobacter mathranii*. Appl. Microbiol. Biotechnol. 64,816–822.

9. S. J. Lee *et al*., (2012). Homologous alkalophilic and acidophilic L-arabinose isomerases reveal region-specific contributions to the pH dependence of activity and

stability. Appl Environ Microbiol 78, 8813-8816.

10. Y. H. Hong, D. W. Lee, Y. R. Pyun, S. H. Lee, (2011). Creation of metal-independent hyperthermophilic L-arabinose isomerase by homologous recombination. J Agric Food Chem 59, 12939-12947.

11. C. Vieille, G. J. Zeikus, (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65, 1-43.

12. Sriprapundh D, Vieille C, Zeikus JG (2003). Directed evolution of *Thermotoga neapolitana* xylose isomerase: high activity on glucose at low temperature and low pH. Protein Eng 16:683–690.

13. Kim BC, Lee YH, Lee HS, Lee DW, Choe EA, Pyun YR (2002). Cloning, expression, and characterization of L-arabinose isomerase from *Thermotoga neapolitana*: bioconversion of D-galactose to D-tagatose using the enzyme. FEMS Microbiol Lett 212:121–126.

14. 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.

15. 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.

16. Tilbeurgh H van, Jenkins J, Chiadmi M, Janin J, Wodak SJ, MrabetNT, Lambeir AM (1992). Protein engineering of xylose (glucose) isomerase from *Actinoplanes missouriensis*. 3.Changing metal specificity and the pH profile by site-directed mutagenesis. Biochemistry 31:5467–5471.

17. Patrick, J. W. & Lee, N. (1969). Subunit structure of L-arabinose isomerase from *Escherichia coli*. J. Biol. Chem.244, 4277–4283.

18. D.-W. Lee *et al*., (2005). A thermodynamic study of mesophilic, thermophilic, and hyperthermophilic l-arabinose isomerases: The effects of divalent metal ions on protein stability at elevated temperatures. FEBS Letters 579, 1261-1266.

19. Wallace, L. J., Eiserling, F. A. & Wilcox, G. (1978). The shape of L-arabinose isomerase from *Escherichia coli.* J.Biol. Chem. 253, 3717–3720.

20. Pauley, J., Power, J. & Irr, J. (1972). L-arabinose isomerase formation in a conditional mutant of gene araA of *Escherichia coli* B-r. J. Bacteriol. 112, 1247–1253.

21. Patrick, J. W., Lee, N., Barnes, N. B. & Englesberg, E. (1971). Coordination of enzyme synthesis in the L-arabinose operon in *Escherichia coli*. I. The effect of manganous ion on the synthesis of L-arabinose isomerase. J. Biol. Chem. 246, 5102–5106.

22. Lee, D. W., Hong, Y. H., Choe, E. A., Lee, S. J., Kim, S. B. Lee, H. S. *et al*. (2005). A thermodynamic study of mesophilic, thermophilic, and hyperthermophilic L-arabinose isomerases: The effects of divalent metal ions on protein stability at elevated temperatures. FEBS Letters, 579, 1261–1266.

23. Izumuri, K., Miyoshi, T., Tokuda, S. & Yamabe, K. (1984). Production of D-tagatose from dulcitol by *Arthrobacter globiformis*. Appl. Environ. Microbiol. 46,1055–1057.

24. R. Schleif (2000). Regulation of the L-arabinose operon of *Escherichia coli*. Trends in genetics: TIG 16, 559-565.

25. J. M. Choi *et al*., (2016). Structure of the thermophilic L-arabinose isomerase from *Geobacillus kaustophilus* reveals metal-mediated intersubunit interactions for activity and thermostability. Arch Biochem Biophys 596, 51-62.

26. D.-W. Lee *et al*., (2005). Distinct metal dependence for catalytic and structural functions in the l-arabinose isomerases from the mesophilic *Bacillus halodurans* and the thermophilic *Geobacillus stearothermophilus*. Archives of Biochemistry and Biophysics

434, 333-343.

27. D. W. Lee *et al*., (2004). Characterization of a thermostable L-arabinose (D-galactose) isomerase from the hyperthermophilic eubacterium *Thermotoga maritima*. Appl Environ Microbiol 70, 1397-1404.

28. J. E. Vick, D. M. Z. Schmidt, J. A. Gerlt (2005). Evolutionary potential of (β/α)8-Barrels: in vitro enhancement of a "new" reaction in the enolase superfamily. Biochemistry 44, 11722-11729.

29. S. J. Lee *et al*., (2005). Characterization of a thermoacidophilic L-arabinose isomerase from *Alicyclobacillus acidocaldarius*: role of Lys-269 in pH optimum. Appl Environ Microbiol 71, 7888-7896.

30. Kim BC, Grote R, Lee DW, Antranikian G, Pyun YR. (2001). *Thermoanaerobacter yonseiensis sp.* nov., a novel extremely thermophilic, xylose utilizing bacterium that grows at up to 85 degrees C. Int. J. Syst. Evol. Microbiol. 51:1539 –1548.

31. Chen Y, *et al*. (2011). Complete genome sequence of *Alicyclobacillus acidocaldarius* strain Tc-4-1. J. Bacteriol. 193:5602–5603.

32. Lee SJ, *et al*. 2005. Characterization of a thermoacidophilic L-arabinose isomerase from *Alicyclobacillus acidocaldarius*: role of Lys-269 in pH optimum. Appl. Environ. Microbiol. 71:7888 –7896.

33. Mavromatis K, *et al*. 2010. Complete genome sequence of *Alicyclobacillus acidocaldarius* type strain (104-IA). Stand. Genomic Sci. 2:9 –18.

34. Lee DW, *et al*. (2005). Distinct metal dependence for catalytic and structural functions in the L-arabinose isomerases from the mesophilic *Bacillus halodurans* and the thermophilic *Geobacillus stearothermophilus.* Arch.Biochem. Biophys. 434:333–343.

35. William Brown, Amy Ralston, Kenna Shaw (2008). Positive Transcription Control: The Glucose Effect. Nature Education.

36. Hichem Chouayekh, Wacim Bejar, Moez Rhimi, Karim Jelleli, Malek Mseddi, Samir Bejar (2007). Characterization of an L-arabinose isomerase from the *Lactobacillus plantarum* NC8 strain showing pronounced stability at acidic pH. FEMS Microbiol Lett 277: 260–267.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest appreciation to my supervisor, Prof. Sung Haeng Lee, Ph.D., who has taught me science, given me inspiration in the research, and advised me to complete this dissertation.

My dissertation would not be completed without the considerations and reviews from the Committee of my dissertation defense.

I would like to thank Prof. Tae-Hyoung Kim, Chairman of the Committee, who has considered and straightly given the critical comments to improve my dissertation.

I would like to thank Prof. Jung-Hee Park, Member of the Committee, who has considered my dissertation defense and given me valuable comments to improve my writing.

I would like to thank the staffs at beamlines PLS-7A at the Pohang Light Source (PLS) for technical support during data collection.

Furthermore, I would like to thank Dr. Immanuel Dhanasingh, my senior, for his teaching and helping when we worked together in our lab and, especially, for his valuable comments on the dissertation. I would like to thank Hye-Jin Kwon, my labmate, who helped me read and translate the abstract in Korean version and prepared my final presentation with her excellent skills.

My family, my mother and my younger brother, is where I can always find my happiness and motivation for all my efforts.