Structure and Biosynthesis of Pellicle Polysaccharides in a Thermotolerant Acetic Acid Bacterium, *Acetobacter tropicalis* SKU1100

（耐熱性酢酸菌 *Acetobacter tropicalis* SKU1100 における菌膜多糖の構造とその生合成）

Ibnaof Ali Ibnaof Ali

Graduate School of Medicine

Yamaguchi University

2011

Graduate School of Medicine

Yamaguchi University

Structure and biosynthesis of pellicle polysaccharides in a thermotolerant acetic acid bacterium, *Acetobacter tropicalis* SKU1100

Ibnaof Ali Ibnaof Ali

Doctor of Applied Molecular Bioscience

Graduate School of Medicine

Yamaguchi University

2011

CONTENTS

LIST of TABLES

Table **page**

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST of FIGURES

Figure **page**

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

LIST of ABBREVIATIONS

AAB	acetic acid bacteria
ADH	alcohol dehydrogenase
Ala	alanine
ALDH	aldehyde dehydrogenase
Arg	arginine
Asn	asparagine
Asp	aspartic acid
C	cytocine
COSY	correlation spectroscopy
CPS	Capsular polysaccharide
CTAB	Cetyltrimethyl ammonium bromide
Cys	cysteine
DMSO	dimethyl sulphoxide
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
EPS	Extracellular polysaccharide
et al	and others
g	gram (s)
Galf	galactofuranose
Galp	galactopyranose
GC-MS	Gas chromatography mass spectrometer
Glcp	glucopyranose
Gln	glutamine
Glu	glutamic
Gly	glycine
h	hour (s)
His	histidine
HPLC	high-performance liquid chromatograpy
HSQC	heteronuclear single quantum coherence
Ile	isoleucine

kb	kilobase (s)
KD	kilodalton (s)
L	litre
LB	Luria-Bertani medium
Leu	leucine
Ly	lysine
Met	methionine
mg	milligram (s)
min	minute (s)
ml	milliliter (s)
mM	millimolar
ng	nanogram (s)
NMR	nuclear magnetic resonance
NOESY	NuclearOverhauserenhancement spectroscopy
°C	Degree Celsius
ORF	open reading frame
PCP	polysaccharide co-polymerase proteins
PCR	Polymerase Chain Reaction
Phe	phenylalanine
PMAA	partially methylated alditol acetates
Pro	proline
R	rough surfaced colony
Rhap	rhamnopyranose
rpm	revolutions per minute
S	smooth surfaced colony
sec	second (s)
Ser	serine
TAIL PCR	Thermal asymmetric interlaced PCR
TFA	trifluoroacetic acid
Thr	threonine
TMS	trimethylsilyl
TOCSY	total correlation spectroscopy
Tris-HCl	Tris-(hydroxymethyl)-aminoethane
Trp	tryptophan

TSP	sodium3-trimethylsilyl-(2,2,3,3,-$^{2}H_4$)-propanoate
Tyr	tyrosine
UGM	UDP-galactopyranose mutase
Val	valine
µg	microgram (s)
µl	microlitre

GENERAL INTRODUCTION

1. Acetic acid bacteria

Acetic acid bacteria (AAB) are Gram-negative, obligate aerobes that traditionally consist of two genera: *Acetobacter*, and *Gluconobacter*, but presently 12 genera are accommodated to family *Acetobacteraceae* in the *Alpha proteobacteria*: *Acetobacter*, *Gluconobacter*, *Gluconacetobacter*, *Acidomonas*, *Asaia*, *Kozakia*, *Swaminathania*, *Saccharibacter*, *Neoasaia*, *Granulibacter*, *Tanticharoenia*, and *Ameyamaea* (Yamada and Yukphan, 2008; Yukphan et al., 2008, 2009). AAB have been isolated from many sources including fruits and flowers, and also spoiled or fermented fruits. Because of their high ability to oxidize alcohol and sugars into organic acids, AAB are used in production of fermented foods and beverages such as vinegar, nata, kombucha, kefir, koumiss and cocoa products (De Vuyst et al., 2008). Moreover, AAB are used in biotechnology for production of L-sorbose and 2-keto-L-gluconic acid, which are intermediates for production of vitamin C, D-tagatose, which is used as bulking agent in foods and noncalorific sweeteners, and shikimate that is a key intermediate in many antibiotics, and also biocellulose (Raspor and Goranovič, 2008). On the other hand, AAB cause spoilage of some foods and beverages and recently *Granulibacter bethesdensis* was found to be pathogenic (Greenberg et al., 2006).

Acetobacter species have high capacity of acetic acid production due to the high aerobic membrane-bound respiration (Matsushita et al., 1982). Quinohemoprotein alcohol dehydrogenase (ADH) and mlybdopterin enzyme aldehyde dehydrogenase (ALDH) are the enzymes of interest. Detailed study has shown that the substrate oxidation by the membrane-bound dehydrogenases is coupled with the respiratory chain of the bacteria and the electrons generated are transferred to the terminal oxidases yielding energy (Matsushita et al., 2002). Two methods are used for vinegar production, traditional static fermentation and submerged fermentation. In the submerged culture, AAB are incubated in the liquid media and oxygen is added continuously, this method has a shorter fermentation time (1-2 days) but with higher costs. While in static culture, fermentation is started by the addition of unidentified starter culture obtained from the previous fermentation called seed-vinegar. After few days, a pellicle of AAB covers the surface of fermentation pond. This technique has lower costs and produce high quality vinegar, though it takes longer time for cultivation.

2. Thermotolerance in Acetic acid bacteria

Thermotolerant bacteria are organisms that can grow at 5-10°C higher than their mesophilic strains in the same species (Manaia and Moore 2002). Usual AAB are mesophilic bacteria with optimum growth temperature of 25-30°C. However, thermotolerant AAB that can grow at 37-40°C were isolated from fruits in Thailand with the purpose of developing vinegar fermentation at higher temperature that may reduce cooling costs (Saeki et al., 1997). Most of the isolates were classified as *Acetobacter pasteurianus*, and *Acetobacter tropicalis* that can produce acetic acid and polysaccharides, respectively, up to 40°C. The thermotolerant *A. pasteurianus* could grow at 4% acetic acid and oxidize up to 9% ethanol without lag phase while their mesophilic counterparts can grow on 2% acetic acid and show delay or no ethanol oxidation in such high concentration. Thermotolerant AAB, namely *A. pasteurianus* and *A. tropicalis,* were also isolated from products in tropical sub-Sahara in Africa (Ndoye et al., 2006).

To clarify the ability of thermotolerant AAB to grow and produce acetic acid at high temperature, ADHs from thermotolerant *A. pasteurianus* MSU10, thermotolerant *A. pasteurianus* SKU1108, and mesophilic *A. pasteurianus* IFO3191 strains were purified and characterized. ADHs of the thermotolerant strains showed higher optimum temperature and more heat stability than one of the mesophilic strains (IFO3191). Moreover, ADHs from MSU10 and SKU1108 displayed higher resistance to ethanol and acetic acid than IFO3191 at elevated temperature. ADH genes were also cloned and their amino acid sequences were found to have some difference between enzymes of MSU10 and SKU1108, and one of IFO3191. The sequence difference could be partially related to the thermotolerance of these thermotolerant strains (Kanchanarach et al., 2010).

3. Pellicle polysaccharides of Acetic acid bacteria

Biofilm formation is an essential survival strategy for microorganisms. The biofilm can be formed on surfaces (biotic or abiotic), and on air-liquid interfaces as pellicle. Matrix production is an important feature of biofilm that provides structural support and facilitates the exchange of small molecules. Polysaccharides, proteins, and DNA are the major components of the matrix (Karatan and Watnick, 2009). Bacterial polysaccharides are found as O-antigen of lipopolysaccharide, capsular polysaccharide (CPS) that is covalently bound to the cell membrane, and exopolysaccharide (EPS) which is not attached to the cell surface and secreted in the growth medium. Polysaccharides are classified into two types depending

on their composition: homopolysaccharides composed of one kind of monosaccharides, and heteropolysaccharides composed of more than one kind of monosaccharide. They are important in bacteria-host interactions (Roberts, 1996); food, pharmaceuticals, and fine chemical industries (Sanford and Baird, 1983).

Acetobacter species are well known to produce pellicle on the surfaces of the growth media in order to keep high aeration state. This pellicle contained bacterial cells attached to each other via CPS and other extracellular matrix. The pellicle CPS is found as homopolysaccharides like cellulose, which is produced by *Gluconacetobacter xylinus,* or as heteropolysaccharide such as the pellicles of *A. pasteurianus* IFO3284 and *A. toropicalis* SKU1100 that are composed of glucose and rhamnose (Moonmangmee et al., 2002a), and of glucose, galactose, and rhamnose (Moonmangmee et al., 2002b), respectively. Many *Acetobacter* strains including *A. pasteurianus* 3284 and *A. tropicalis* SKU1100 are able to form two types of colony on agar surface: rough surfaced colony (R), and smooth surfaced colony (S). The R strain can produce pellicle which allows it to float on the medium surface in static culture, while S strain cannot produce pellicle in static culture. The R and S phenotypes are exchangeable by spontaneous mutation (Matsushita et al., 1992).

Genes involved in biosynthesis of cellulose and acetan in *Ga. xylinus* have been explored (Wong et al., 1990; Saxena et al., 1994; Ishida et al., 2002). Moreover, a gene operon, *polABCDE*, involved in pellicle formation in the R strain of *A. tropicalis* SKU1100 has been identified (Deeraksa et al., 2005). In this operon, *polABCD*, exhibits high similarity to *rfbBACD* genes which are involved in dTDP-L-rhamnose biosynthesis, while *polE*, a novel gene downstream of *polABCD,* has a relatively low similarity to glycosyltransferases. The disruption of this gene caused the bacteria not to form pellicle in static culture because of no capsular polysaccharide production. Interestingly instead, the mutant cells secreted EPS in the culture medium (an S strain like phenotype). Moreover, the mutation sites in S strain of *A. tropicalis* SKU1100 was shown to be in a 7 C repeated residues in the coding region of *polE* gene. Hence, it was hypothesized that the *polE* gene may be involved in the switching of CPS to EPS. In addition, *galE*, a gene that encodes UDP-galactose 4-epimerase involved in UDP-galactose biosynthesis, has also been identified in *A. tropicalis* SKU1100. The disruption of this gene also resulted in no pellicle formation in static culture but secretion of EPS composed of glucose and rhamnose in culture medium (Deeraksa et al., 2006).

Structural characterization of polysaccharides is important to understand their biosynthesis, biological functions, and their physical properties. However, it is not so easy to know the structure, because there is no simple method to manipulate them like PCR for DNA. To analyze the structure of bacterial polysaccharides, thus they must be first separated from cells or culture media and purified by column chromatography. Then their monosaccharide composition should be known using HPLC or GC-MS after derivatization. Thereafter, the linkage positions of these monosaccharides could be known using methylation and GC-MS analysis. In addition, NMR analysis will confirm the chemical analysis and give information about repeating unit size and monosaccharide sequence in the repeating units, as well as the presence of non-carbohydrate groups like methyl, acetyl, sulphate, and phosphate. NMR methods of interest are: 1D ^{1}H and ^{13}C which give information on repeat unit size; 2D COSY, TOCSY, and ^{1}H-^{13}C HSQC which give information on monosaccharide identity by determining the chemical shift values of protons and carbons. Finally sequence information could be obtained from 2D NOESY or HMBC experiments (Duus et al., 2000).

Hexose sugars are usually found in nature in thermodynamically favorable pyranose form. Nevertheless, galactofuranose (Galf) present in glycoconjugates of bacteria, fungi, and parasites. Galf is formed by enzymatic reaction of UDP-galactopyranose mutase (UGM). This enzyme is encoded by *glf* gene which has been identified in *Escherichia coli* (Nassau et al., 1996), *Klebsiella pneumonia* (Köplin et al., 1997), *Mycobacterium tuberculosis* (Weston et al., 1998) and many eukaryotic parasites (Beverley et al., 2005). Moreover, a number of putative UGMs have been identified in many microbial species. Galf residues are essential for cell viability or play important role in cell physiology because Galf containing surface glycoconjugate has been shown to be important for viability or pathogenecity of some microorganism. For instance, in *Mycobacterium smegmatis, glf* gene is essential for cell viability because it involves in arabinogalactan polysaccharide biosynthesis. The arabinogalactan attaches the mycolic acid layer to the peptidoglycan layer, and the *glf* gene could be deleted only in the presence of two rescue plasmids containing *glf* gene and galactofuranosyltransferase (Pan, et al., 2001). Due to its absence in mammalian cells, *glf* is an attractive target for antimicrobial agents development (Pedersen and Turco, 2003)

OBJECTIVES

A.tropicalis SKU1100 is a thermotolerant AAB capable of producing pellicle at higher temperature. Genetic investigation revealed a gene cluster containing a novel gene, *polE*, and the spontaneous mutation or disruption of this gene resulted in no pellicle formation, instead, production of EPS with higher molecular weight. Likewise, the disruption of *galE* gene resulted in the same phenotype and production of EPS with higher molecular weight. Moreover, draft genome sequence of *A. tropicalis* is now available and thus it can be compared with the complete genome sequence of mesophilic *A. pasteurianus*. Therefore, the objectives of this study are:

1. To elucidate the structures of *A.tropicalis* SKU1100 wild type CPS and of Δ*polE* and Δ*galE* EPSs to gain insight in PolE function and polysaccharide biosynthesis.

2. To compare the amino acid substitutions of orthologous proteins between *A. tropicalis* and mesophilic *A. pasteuranius* IFO3283 to understand the background of thermotolerance in *A. tropicalis* SKU1100.

3. To search *glf* gene and to explore galactofuranose biosynthesis in *A. tropicalis* SKU1100.

CHAPTER 1

Increased number of Arginine-based salt bridges contributes to the thermotolerance of thermotolerant acetic acid bacteria, *Acetobacter tropicalis* SKU1100

ABSTRACT

Thermotolerant acetic acid bacteria (AAB), *Acetobacter tropicalis* SKU1100, can grow above 40 °C. To investigate the basis of its thermotolerance, we compared the genome of *A. tropicalis* SKU1100 with that of mesophilic AAB strain *Acetobacter pasteurianus* IFO3283-01. The comparative genomic study showed that amino acid substitutions from large to small residue and Lys to Arg occur in many orthologous genes. Furthermore, comparative modeling study was carried out with the orthologous proteins between SKU1100 and IFO3283-01 strains, indicating that the number of Arg-based salt bridges increased in protein models. Since it has been reported that Arg-based salt bridges are important factor for thermo-stability of protein structure, our results strongly suggest that the increased number of Arg-based salt bridges may contributes to the thermotolerance of *A. tropicalis* SKU1100 (the thermo-stability of proteins in *A. tropicalis* SKU1100).

INTRODUCTION

Acetic acid bacteria (AAB) are Gram-negative strictly aerobic bacteria, which are classified into twelve genera, of which the major genera are *Acetobacter, Gluconobacter,* and *Gluconacetobacter* (Cleenwerck et al., 2002; Loganathan and Nair, 2004; Jojima et al., 2004; Yukphan et al., 2005, 2008, 2009; Greenberg et al., 2006). AAB oxidize various sugars and alcohols into the corresponding acids. In particular, *Acetobacter* sp. has been used from ancient times for industrial vinegar fermentation because of their high ethanol-oxidizing ability. Industrial acetic acid fermentation is carried out at 25-30 °C, and thus requires a cooling system to maintain the temperature of the culture from the fermentative heat generation. Recent global warming leads to further expenses for the cooling. Therefore, favorable AAB that can work properly near or above 40 °C are needed.

Thermotolerant bacteria are so-called mesophilic bacteria but grow at temperatures 5 to 10 °C higher than typical mesophilic strains of the same genus or the same species (Manaia and Moore, 2002; Ndoye et al., 2006). These strains are different from so-called thermophilic bacteria, which are defined as strains with an innate ability to grow above 60°C (Saeki et al., 1997; Sikorski et al., 2008). Therefore, it is conceivable that these thermotolerant strains have acquired their growth phenotype by adapting to a habitat with a higher temperature, such as tropical regions (Ndoye et al., 2006; Saeki et al., 1997).

Previously, we isolated a thermotolerant strain *Acetobacter* sp. SKU1100 from fruits in Thailand, which is now called *Acetobacter tropicalis* SKU1100 (Deeraksa et al., 2005, 2006; Moonmangmee et al., 2002). This strain can grow in shaking culture above 40 °C, which is higher than that of *Acetobacter pasteurianus* NBRC3283, which can grow up to 38 °C (non-acetic acid fermentation condition). In addition to the complete genome of *A. pasteurianus* NBRC3283-01(Azuma et al., 2009), we have recently obtained the draft genome of *A. tropicalis* SKU1100 (GenBank accession: BABS01000001-BABS01000773), which may enable us to elucidate the thermotolerance mechanism of AAB. It has been reported that Lys to Arg substitutions are the important factor to generate thermo-stability in protein structure of thermophilic bacteria (Vieille and Zeikus, 2001). These substitutions are also thought to be the important factor to generate thermo-stability in protein structure of thermotolerant bacteria (Nishio et al., 2003). In addition, it has been suggested that salt bridges formed with Arg residues, rather than Lys residues, are more stable and thus contribute to the thermo-stability of proteins (Bae and Phillips, 2004, 2005; Liu et al., 2008).

In the present study, to investigate the relationship between thermotolerance and amino acid mutation of orthologous proteins for AAB, *A. tropicalis* SKU1100 genome was compared with that of mesophilic AAB strain *A. pasteurianus* IFO3283-01 in terms of amino acid substitutions. In addition, 3-D models of the orthologous proteins were constructed to elucidate salt bridge-forming residues, because of Lys to Arg substitutions being increased in the thermotolerant strain. The results obtained in our study clearly suggested that Arg-based salt bridge have contributed to evolutional change in this strain in terms of its thermotolerance.

MATERIALS and METHODS

Bacterial strains

A. tropicalis SKU1100 (NBRC 101654) and *A. pasteurianus* NBRC3283 were used in this study. These strains were cultured in potato medium (0.5% glucose [D], 2% glycerol [G], 1% yeast extract [Y], 1% peptone [P], and 150 ml/L of potato extract) or YPGD medium (0.5% each of Y, P, G, and D) (Deeraksa et al., 2005).

Comparative analysis of orthologous genes

The draft genome assembly (GenBank accession numbers: BABS01000001-BABS01000773) of *A. tropicalis* SKU1100 (NBRC 101654) and the previously published complete genome sequence of *A. pasteurianus* IFO3283-01 were used for the comparative genome analysis (Azuma et al., 2009). Although *A. pasteurianus* IFO3283-01 have 6 plasmid sequences, only the chromosome sequence was used for this analysis. Homologous genes were identified by homology searches in an amino acid sequence using the BLASTP filtering expectation value of identity $\geq$ 70%, e-value $\leq$ 1.0E-30, and sequence overlap $\geq$ 70% (Altschul et al., 1997). When the best hit was identical to the query, the ORFs were regarded as being orthologous genes. Putative localization of orthologous proteins was evaluated using the PSORTb program (Gardy, Spencer, et al., 2003; Gardy, Laird, et al 2005). Multiple sequence alignment between orthologous proteins was performed using ClustalW (Larkin et al., 2007). Nucleotide sequence alignment of orthologous genes was achieved from amino acid sequence alignment using house-written ruby script.

Comparative modeling of orthologous proteins

Initially, BLASTP search of *A. pasteurianus* orthologous proteins were performed against pdbaa database (identity ≥ 20%, e-value ≤ 1.0 E-10, and sequence overlap ≥ 70%) to find homologous sequences of known protein structures. The pdbaa database was downloaded from the NCBI FTP site at ftp://ftp.ncbi.nih.gov/blast/db/. Satisfying the BLASTP criteria, sequence alignments between query and top hit protein were constructed using ClustalW (Larkin et al., 2007). From sequence alignment, 3-D models of *A. pasteurianus* orthologous proteins were built using Modeller 9v7 (Eswar et al., 2006; Martí-Renom et al., 2000; Sali and Blundell, 1993; Fiser et al., 2000). In addition, 3-D models of *A. tropicalis* were constructed with *A. pasteurianus* 3-D model structure as template. The resulting 3-D models were used for the detection of salt bridge interaction. The criterion for determining salt bridges was as follows: distance between carboxyl oxygen atoms on the side chain of Glu or Asp and ε-nitrogen atoms on the side chain of Arg within 5.00 Å (Liu et al., 2008).

RESULTS and DISCUSSION

Characterization of *A. tropicalis* SKU1100 and *A. pasteurianus* NBRC3283

The growth abilities of the two closely related species, *A. tropicalis* SKU1100 and *A. pasteurianus* NBRC3283, were compared at high temperatures. As shown in Fig.1.1, it was found that both species could grow well on YPGD agar plate up to 39°C, although SKU1100 grew better than NBRC3283. However, over 40°C, the growth of NBRC3283 was largely repressed, while SKU1100 could grow even at 41°C. Thus, *A. tropicalis* SKU1100 was considered as a thermotolerant AAB strain than *A. pasteurianus* NBRC 3283.

Comparison of amino acid composition of orthologous groups

The 1,242 highly conserved orthologous groups between *A. tropicalis* and *A. pasteurianus* were determined (see Materials and Methods) and used for the comparison of amino acid composition. The number of orthologous genes as a function of GC content is shown in Fig.1.2. The peak of ORF number in *A. tropicalis* shifts to a higher GC content than that of *A. pasteurianus*. The differences in average GC content between the two microorganisms are directly reflected in the GC content of the ORFs. The amino acid composition of the protein coding regions was analyzed, as shown in Table 1.1. Ala, Arg, Asp, Gly, Leu, Thr, and Val are more frequently used in *A. tropicalis* than *A. pasteurianus*, which results in the higher GC content of *A. tropicalis*.

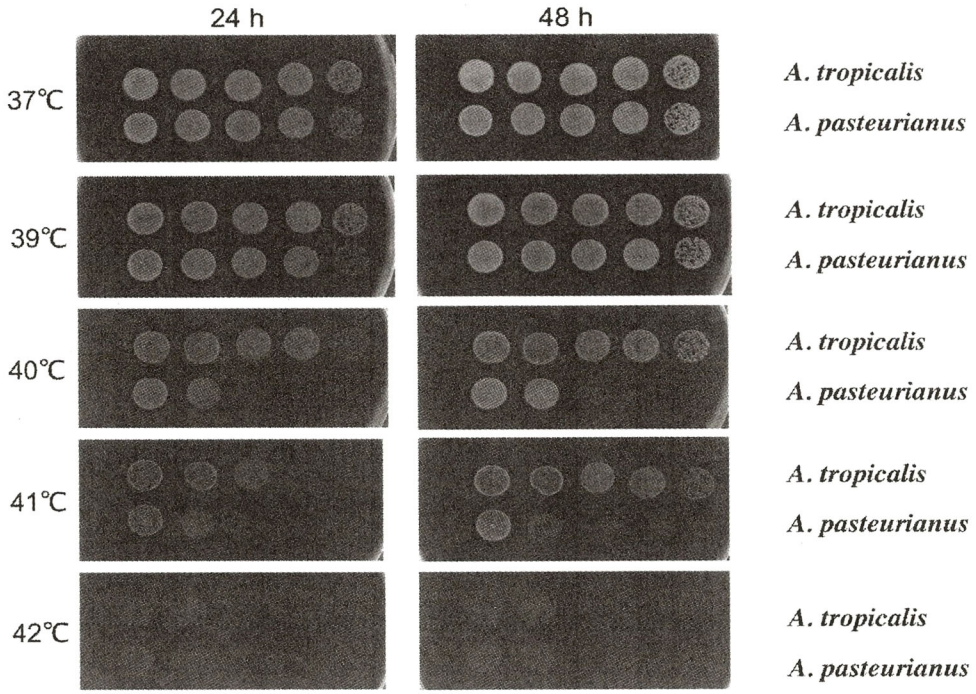


Fig. 1.1 Growth comparison of *A. tropicalis* SKU1100 and *A. pasteurianus* NBRC3283 at various temperatures under non-fermentation condition.

These strains were pre-cultured in potato medium until its turbidity reached 150 Klett units at 30 °C. Then, the cultures were diluted 10^{-1}, 10^{-2}, 10^{-3}, or 10^{-4} times. 7 μl of the diluted solutions was spotted onto YPGD plates, which were incubated at 37 °C, 39 °C, 40 °C, 41 °C, or 42 °C for 24 h and 48 h. In each panel, the upper lane shows the spots of *A. tropicalis* and the lower lane shows those of *A. pasteurianus*.

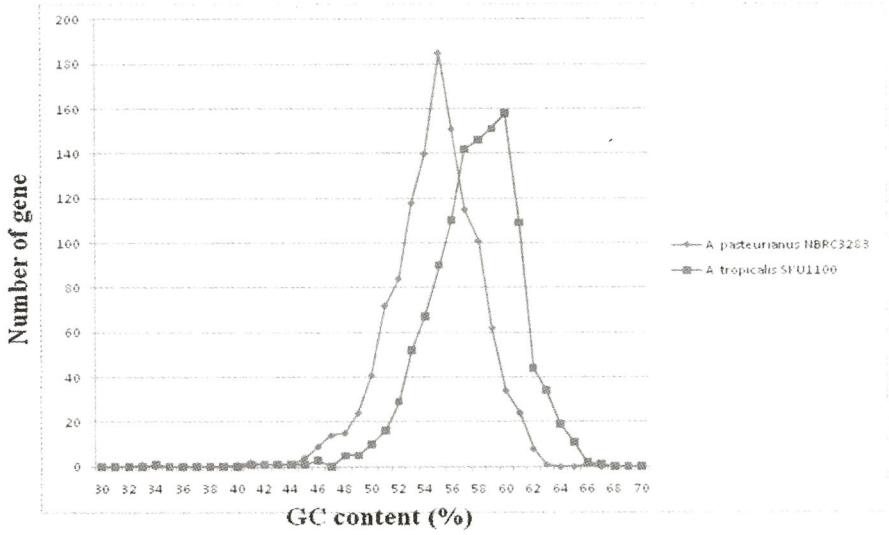


Fig. 1.2 Distributio of GC content of 1,242 orthologous groups between *A. pasterianus* NBRC3283 and A. *tropicalis* SKU1100.

The number of codons per 1,000 bases (fraction values) of GCC (Ala), GCG (Ala), GAC (Asp), GGA (Gly), GGC (Gly), GGG (Gly), CTC (Leu), CTG (Leu), CGG (Arg), Thr (ACC), and Thr (ACG) are increased (data not shown). Nishio and coworkers (2003) have reported that thermotolerant corynebacterium, *Corynebacterium efficiens*, has more rare codons, GGG (Gly) and CGG (Arg), than mesophilic *Corynebacterium glutamicum*. Our results showed the same tendency with respect to CGG (Arg) codon. Increased number of CGG codons may be correlated with thermotolerance of bacteria, as described below.

Table 1.1 Amino acid composition of protein coding regions

	Number		Percentage (%)	
Amino acid	*A. pasteurianus*	*A. tropicalis*	*A. pasteurianus*	*A. tropicalis*
Ala	52,531	53,062	11.55	11.66
Arg	30,006	30,584	6.60	6.72
Asn	13,086	12,292	2.88	2.70
Asp	24,580	25,095	5.40	5.52
Cys	4,565	4,393	1.00	0.97
Gln	16,433	15,724	3.61	3.46
Glu	26,011	25,716	5.72	5.65
Gly	38,044	38,690	8.36	8.50
His	10,739	10,504	2.36	2.31
Ile	23,081	22,544	5.07	4.96
Leu	46,125	46,647	10.14	10.25
Lys	16,680	16,109	3.67	3.54
Met	11,880	11,506	2.61	2.53
Phe	16,331	16,258	3.59	3.57
Pro	24,569	24,745	5.40	5.44
Ser	24,504	24,815	5.39	5.45
Thr	25,525	25,927	5.61	5.70
Trp	5,913	5,867	1.30	1.29
Tyr	10,362	10,278	2.28	2.26
Val	33,984	34,193	7.47	7.52
Total	454,949	454,949	-	-

Comparison of amino acid substitution of orthologous groups

To investigate the direction of amino acid mutation, the biased amino acid substitution of 1,242 orthologous genes between *A. pasteurianus* and *A. tropicalis* was calculated. The amino acid sequences of 1,242 orthologous proteins were aligned using ClustalW (Larkin et al., 2007). Amino acid substitutions between *A. pasteurianus* and *A. tropicalis* were extracted and compared (Table 1.2).

The results clearly showed high substitution frequency of large to small residues (Met-Leu, Glu-Asp, Lys-Arg, Glu-Ala and Ile-Val) from *A. pasteurianus* to *A. tropicalis*. It has been reported that large to small hydrophobic residue substitutions of subunit-subunit interface stabilize the thermal stability of protein (Ohkuri and Yamagishi, 2003). However, some previous studies have suggested that large to small hydrophobic residue substitutions of buried hydrophobic core destabilize the thermal stability of protein (Vieille and Zeikus, 2001; Pace, 1992). Thus, although such many amino acid substitutions occur between each two orthologous proteins, it is difficult to evaluate the effect of each amino acid substitution. Even though, these mutations can be speculated to make more compact packing structure, and to decrease the conformational flexibility. The decreasing of conformational flexibility is important factor for the thermal stability of protein (Liu et al., 2008). Therefore, we hypnotized that the decreasing of conformational flexibility due to such the amino acid exchange stabilizes the protein structure in thermotolerant strain. Of the amino acid substitutions occurred, Lys to Arg substitutions from *A. pasteurianus* to *A. tropicalis* were paid attention. Some previous studies have suggested that Lys to Arg substitutions affect the thermal stability of proteins in both thermophilic and thermotolerant bacteria (Vieille and Zeikus, 2001; Nishio et al., 2003). Several properties of Arg residues are thought to be better adapted to high temperatures than Lys residues (Vieille and Zeikus, 2001), including the fact that the Arg-δ-guanido moiety has a reduced chemical reactivity due to its high pKa and its resonance stabilization. Indeed, it has been reported that thermotolerant *C. efficiens* have more Arg residues than mesophilic *C. glutamicum* (Nishio et al., 2003).

In addition, Arg or Lys residue has been known to form a salt bridge with either Glu or Asp residue, and it has been suggested that the salt bridge formed with Arg and Glu (or Asp) confers a strong stabilizing mechanism for hyperthermophilic proteins (Vieille and Zeikus, 2001). Besides, many other substitutions to Asp and Arg were observed (Asn-Asp,

Gln-Arg and His-Arg). Increasing of Asp and Arg residue may contribute to the increasing of the Arg-based salt bridges, which contribute to the thermo-stability of *A. tropicalis* proteins.

Table 1.2 Comparison of amino acid substitution patterns in orthologous genes between *A. pasteurianus* IFO3283-01 and *A. tropicalis* SKU1100.

A. pasteurianus	*A. tropicalis*	Forward	Reverse
Glu	Asp	1,889	1,621
Lys	Arg	1,477	1,227
Met	Leu	1,077	842
Asn	Asp	781	548
Ile	Leu	1,310	1,101
Ile	Val	2,535	2,356
Gln	Arg	738	574
Glu	Ala	1,027	872
Asn	Thr	375	235
His	Arg	513	382

Evaluation of the number of salt bridge-forming Arg residues

To evaluate the number of salt bridge-forming Arg residues, comparative models of orthologous proteins were constructed. Comparative modeling predicts the three-dimensional structure of a given target protein sequence primarily on the basis of its alignment to proteins of known structure as templates. The prediction process consists of fold assignment, target-template alignment, and model building. Initially, we performed a BLAST search against the pdbaa database (identity $\geq$ 20%, e-value $\leq$ 1.0E-10, and sequence overlap $\geq$ 70%) to find sequences homologous to proteins of known three-dimensional structures. Satisfying the BLAST criteria, 737 orthologous groups were found to have close homologues in Protein Data Bank (data not shown). Using ClustalW, 737 orthologous amino acid sequences of *A. pasteurianus* were aligned with template sequences. Model building was performed using Modeller 9v7. To reduce alignment error between two orthologous proteins, 3-D models of *A. tropicalis* were constructed with *A. pasteurianus* 3-D model structure as a template. We found 253 substitutions from Lys to Arg in *A. tropicalis* in the 737 orthologous proteins. Fig. 1.3 shows one example of a substitution in the structural model between orthologous NAD-dependent alcohol dehydrogenases of *A. pasteurianus* and *A. tropicalis*. In case of the latter enzyme, Arg 28, which substitutes for Lys of the protein in *A. pasteurianus*, forms a salt

bridge with two Asp residues, Asp 30 and Asp 131. In contrast, original Lys 28 in *A. pasteurianus* has forms a Lys-based salt bridge with only Asp 131 residue.

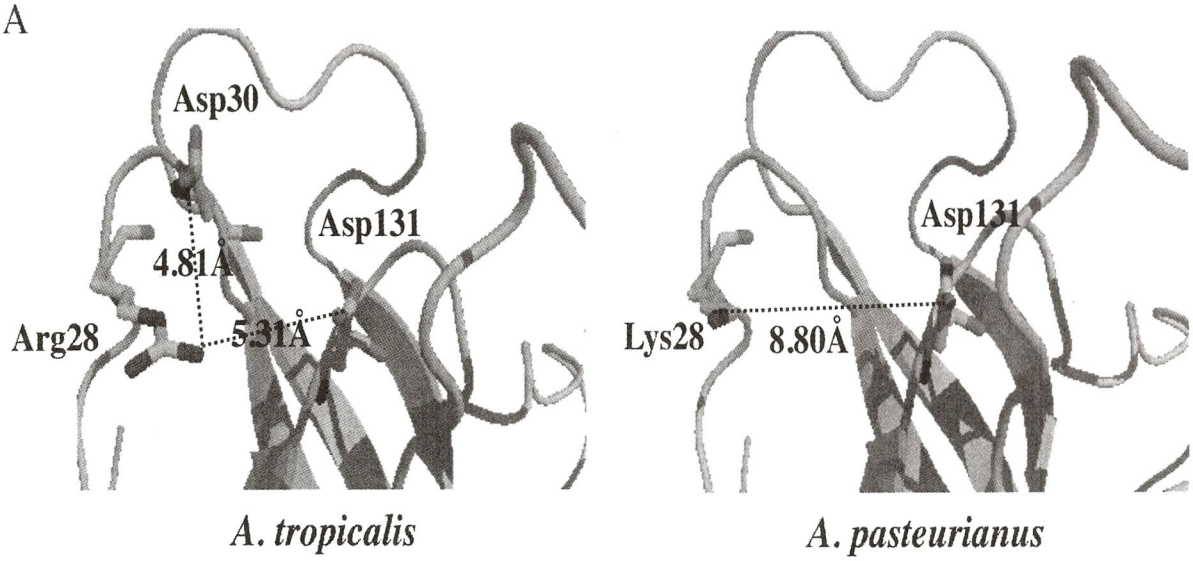
Thus, the total numbers of Arg-based salt bridges were compared in the 737 proteins between *A. tropicalis* and *A. pasteurianus*. As shown in Table 1.3, of the 737 proteins, 363 proteins have a higher number of Arg-based salt bridges in *A. tropicalis* when compared with those of *A. pasteurianus*, while 241 proteins have a lower number of salt bridges in *A. tropicalis*. Another 134 proteins have the same number of residues between *A. tropicalis* and *A. pasteurianus*. When examining the subcellular localization of these proteins, the proteins with a higher number of Arg-based salt bridges in *A. tropicalis* are observed among the soluble proteins, cytoplasmic and periplasmic proteins, while the proteins with less salt bridges are membrane proteins.

Table 1.3 Comparison of the number of salt bridge-forming Arg residues in the orthologous genes.

Three dimensional structure models were constructed for 737 orthologous pairs between *A. tropicalis* and *A. pasteurianus*. In each structural pair, number of Arg-based salt bridges was compared. If the *A. tropicalis* ortholog has higher salt bridge number than *A. pasteurianus* ortholog, it was counted as "Higher in *A. tropicalis*", and vice versa as "Higher in *A. pasteurianus*". Furthermore, these comparisons were separated in their localization.

Subcellular locarization	Total number of orthologous groups	Number of salt bridge-forming Arg residues		
		Higher in *A. tropicalis*	Higher in *A. pasteurianus*	No difference
Cytoplasmic	394	205	123	66
Cytoplasmic Membrane	57	21	28	8
Outer Membrane	13	4	5	4
Periplasmic	17	11	3	3
Extracellular	1	0	1	0
Unknown	255	121	81	53
Total	737	362	241	134

A



B

```
A. pasteurianus      1 -MSGKMKAAVVHEFGKPLTIEELDIPPIKPTQILVKMIACGVCHTDLHAASGDWPKKPHL 59
A. tropicalis        1 -MSGKMKAAVVHEFGKPLTIEELDIPDIRPDQILVKMTACGVCHTDLHAASGDWPKKPTV 59
Template (1LLU_A)    2 TLPQTMKAAVVHAYGAPLRIEEVKVPLPGPGQVLVKIEASGVCHTDLHAAEGDWPVKPPL 61

A. pasteurianus     60 PFIPGHEGVGTVVQVGSEVDWVKEGDVVGVPWLYSACGHCEHCLAGWETLCAKQEETGYS 119
A. tropicalis       60 PFIPGHEGVGVVTKVGSNVTWVKEGDEVGVPWLYSACGHCEHCLGGWETLCAKQQETGYS 119
Template (1LLU_A)   62 PFIPGHEGVGYVAAVGSGVTRVKEGDRVGIPWLYTACGCCEHCLTGWETLCESQQNTGYS 121

A. pasteurianus    120 VNGCFAEYVVADPNYIAHLPKGVDPVKVAPVLCAGLTVYKGLKMTDTRAGNWVAISGVGG 179
A. tropicalis      120 VNGCFAEYVVADPNYVARLPKGIDPVKVAPVLCAGLTVYKGLKMTDTKAGNWVAISGVGG 179
Template (1LLU_A) 122 VNGGYAEYVLADPNYVGILPKNVEFAEIAPILCAGVTVYKGLKQTNARPGQWVAISGIGG 181

A. pasteurianus    180 LGQMAVQYAVAMGLNVVAVDIDDEKLATAKKLGATYTVNARNTDPAAFMQEKVGGVHGGL 239
A. tropicalis      180 LGQMAVQYGVAMGLNVIAVDIDDQKLATARKLGAALTVNARTTDPAAFIQKEVGGVHGGL 239
Template (1LLU_A) 182 LGHVAVQYARAMGLHVAAIDIDDAKLELARKLGASLTVNARQEDPVEAIQRDIGGAHGVL 241

A. pasteurianus    240 ITAVSTKAFSQAMGYVRAGGTLVLNGLPPGDFPISIFDMVMNAITIRGSIVGTRLDMIEA 299
A. tropicalis      240 ITAVSTKAFSQAMGYVRAGGTLVLNGLPPGDFPISIFDMVMKGTTIRGSIVGTRLDMIEA 299
Template (1LLU_A) 242 VTAVSNSAFGQAIGMARRGGTIALVGLPPGDFPTPIFDVVLKGLHIAGSIVGTRADLQEA 301

A. pasteurianus    300 LSFFAEGKVTSVTTTDRIDNINAIFDALKNGRVEGRVVLDFRN 342
A. tropicalis      300 LSFFAEGKVTSVTQTDRLENINNIFTRLEEGKVEGRIVIDFRA 342
Template (1LLU_A) 302 LDFAGEGLVKATIHPGKLDDINQILDQMRAGQIEGRIVLEM-- 342
```

Fig. 1.3 3D modeled structure and multiple alignments with template sequence.

(A) Structural model of NAD-dependent alcohol dehydrogenase (*A. tropicalis* SKU1100: ATPR_0541 and *A. pasteurianus* IFO3283-01: APA01_00250). Salt bridge-forming residues are represented by stick model. In *A. tropicalis* modeled structure, Arg 28 forms salt bridges with two Asp residues, Asp 30 and Asp 131. In contrast, Lys 28 in *A. pasteurianus* modeled structure has forms a Lys-based salt bridge with Asp 131 residue. The figure was prepared using PyMOL 0.99 (Delano, 2002). (B) Salt bridge-forming residue pairs are underlined.

Due to the alignment error and incorrect side chain rotamers, modeling structure is not very accurate in some cases. However, more broad range cut-off value of salt bridges (6-7Å) also showed same tendency (data not shown). Therefore, it had been thought that the increased number of Arg-based salt bridges contributes to the thermo-stability of proteins. Side chain structure of protein sometime has high flexibility in solution. Therefore, to evaluate the more precious Arg-based salt bridge interactions, molecular dynamics (MD) simulations of orthologous proteins should be performed.

In the present study, we investigated the amino acid substitutions between *Acetobacter tropicalis* and *Acetobacter pasteurianus*. Analysis of amino acid substitutions from *A. pasteurianus* to *A. tropicalis* showed that a large number of large to small residue substitutions and Lys to Arg substitutions has occurred, both of which may contribute to increase the thermal stability of proteins. Comparative modeling study of the orthologous proteins showed that the increased number of Arg in *A. tropicalis* orthlogous proteins is largely involved in a salt bridge formation. Recently, Kanchanarach *et al.* (2010) have reported that quinoprotein alcohol dehydrogenases of thermotolerant *Acetobacter* strains have slightly higher optimal temperature and heat stability than those of mesophilic strains. Therefore, we suggest that at least the increased number of Arg-based salt bridges may contribute to the thermo-stability of proteins and thus to the thermotolerance of *A. tropicalis* SKU1100.

CHAPTER 2

Structural characterization of pellicle polysaccharides of *Acetobacter tropicalis* SKU1100 wild type and Δ*polE* mutant strains

ABSTRACT

Mutant of *Acetobacter tropicalis* SKU1100 (R) strain, Δ*polE* defective in pellicle formation, excrete exopolysaccharide (EPS) instead of capsular polysaccharide (CPS) that is produced by the wild-type. We carried out structural analysis of wild type CPS and Δ*polE* mutant EPS using monosaccharide composition analysis, methylation analysis, and ^{1}H and ^{13}C NMR spectroscopy. Wild-type CPS and Δ*polE* mutant EPS had a branched hexasaccharide repeating unit composed of two moles of 2,3-α-L-rhamnopyranosyl, and one mole each of 6-β-D-galactopyranosyl and 2-α-D-glucopyranosyl residues, of which the rhamnosyl residues were branched by terminal-β-D-galactofuranosyl and terminal-α-D-glucopyranosyl residues. By comparing the two structures, it was suggested that PolE may control the switching of EPS to CPS by adding some residue, e.g. β-D-galactopyranosyl residue, to 2,3-α-L-rhamnopyranosyl residue to make 2,3,4-α-L-rhamnosyl residue which leads to CPS formation.

INTRODUCTION

Polysaccharide production is common among both Gram-positive and Gram-negative bacteria. These polysaccharides are classified according to their cellular association into capsular polysaccharide (CPS), which is permanently attached to outer surface of the cells, and exopolysaccharide (EPS) which is secreted into the growth medium. Attention has been paid to bacterial polysaccharides due to their importance in bacteria-host interaction, and biofilm formation (Roberts, 1996), stress adaptation (Ferreira et al., 2010), resistance to desiccation (Ophir & Guntnick, 1994), and their applications in food industry (Sutherland, 1994).

Acetic acid bacteria are Gram-negative obligate aerobes belonging to α-proteobacteria subdivision and well known as vinegar producers. In order to keep high aeration state, almost all *Acetobacter* species have ability to grow floating in static culture by producing a pellicle in the surface of culture medium. This pellicle is an aggregation of cells in the liquid-air interface in which cells are tightly associated with each other by polysaccharides and other extracellular matrix on the cell surface. The pellicle polysaccharides occur as homopolysaccharide of cellulose which is produced by *Gluconacetobacter xylinus*, or as heteropolysaccharides such as CPS produced by *Acetobacter tropicalis* SKU1100 consisting of glucose, galactose, and rhamnose (Moonmangmee et al., 2002b) or CPS of *Acetobacter aceti* IFO3284 (reclassified as *Acetobacter pasteurianus* subsp. Lovaniensis) consisting of glucose and rhamnose (Moonmangmee et al., 2002a). *Acetobacter pasteurianus* IFO3284 produces two different types of colony on agar surface; rough surfaced colony (R strain) and smooth surfaced colony (S strain). The R strain can produce pellicle which allows it to float on the medium surface in static culture, while S strain cannot produce pellicle in static culture. The R and S strains are interconvertible by spontaneous mutation (Matsushita et al., 1992).

The genetic study of polysaccharides in acetic acid bacteria has shown that the *acs* (Saxena et al., 1994) and the *bcs* (Wong et al., 1990) operons, in addition to ORF2 gene (Nakai et al., 2002), are involved in cellulose biosynthesis, and the *aceRQP* operon in acetan biosynthesis in *Ga. xylinus* (Ishida et al., 2002). Moreover, a gene cluster, *polABCDE*, required for pellicle formation in the R strain of *A. tropicalis* SKU1100, has been identified (Deeraksa et al., 2005). In this operon, *polABCD* showed high similarity to *rfbBACD* genes which are involved in d-TDP-L-rhamnose biosynthesis, while *polE*, a novel gene downstream of *polABCD,* had a relatively low similarity to glycosyltransferases. The disruption of this

gene caused the bacteria not to form pellicle in static culture because of no capsular polysaccharide production. Interestingly instead, the mutant cells secreted EPS in the culture medium (an S strain like phenotype). Moreover, the mutation sites in S strain of *A. tropicalis* SKU1100 are found to be in a 7 C repeated residues in the coding region of *polE* gene. Hence, it was hypothesized that the *polE* gene may be involved in the switching of CPS to EPS.

In this study, we elucidated the structures of the wild type CPS as well as EPS of Δ*polE* mutant of *A. tropicalis* SKU1100 (R). Based on the structural information, the function of *polE* gene was also discussed here.

MATERIALS and METHODS

Bacterial strains and culture conditions

A. tropicalis SKU1100 (R) strain and its mutant Δ*polE* (Deeraksa et al., 2005) were used in this study. In Δ*polE* mutant, the *polE* gene was disrupted by insertion of Tn10 transposon. All strains were grown at 30°C in potato medium (Deeraksa et al., 2005). Tetracycline was added at the concentration of 12.5 μg/ml.

Purification of *A. tropicalis* SKU1100 CPS

The CPS was purified according to the method of Moonmangmee et al. (2002b). Briefly, 10% of the seed culture was inoculated to 1 L of potato medium and incubated at 30°C with shaking for 30 h. The cells were then collected, washed 2 times with 50 mM phosphate buffer (pH 6.5), and resuspended in the same buffer (1 g cells/10 ml buffer). The suspension was ultra-sonicated using sonicating microprobe (TP-040, 3mm diameter, TOMY TECH, INC) for 20 min, centrifuged, and ultra-centrifuged to remove cell debris. DNase (50μg/ml) was added to the supernatant and it was incubated at 37°C overnight, followed by an additional overnight incubation with 100μg/ml proteinase K at 37°C. The suspension was then subjected to dialysis against 25 mM Tris-HCl buffer (pH 8.5) overnight. After centrifugation to remove precipitate, the supernatant was applied to a DEAE-cellulose column and eluted with 25 mM Tris-HCl (pH 8.5). Polysaccharide fractions were determined by phenol-sulfuric acid assay (Dubois et al., 1956), pooled, ultra-centrifuged, and precipitated with 2 volumes of cold ethanol. The precipitated polysaccharide was then dissolved in 0.1M NaCl and applied to a Superdex S-200 column. The polysaccharide fractions were pooled and precipitated with 2 volumes of cold ethanol.

Purification of Δ*polE* mutant EPS

The EPSs of Δ*polE* mutant was purified from culture media by basically the same method as described (Deeraksa et al., 2006). The cultivations were performed in 1 L of YPG medium (5% glycerol, 0.5% peptone, and 0.5% yeast extract) for 2 days. The cells were removed by centrifugation (9000 x *g* for 10 min at 4°C). The culture media was then collected and concentrated to one third using ultra filtration (20KD cutoff, Advantec), followed by DEAE-cellulose column chromatography as described above. Polysaccharide fractions were then pooled, and treated as described above, and applied to a Sephacryl S-400 column equilibrated with 0.1 M NaCl. Polysaccharide fractions were combined and precipitated with 2 volumes of cold ethanol.

Monosaccharide composition and molecular size analysis

The monosaccharide composition was analyzed using trimethylsilyl (TMS) methylglycoside method as described by Wozniak et al. (2003). Aliquots of 300-μg of each polysaccharide were lyophilized, and then, mixed with 500 μl of 1M methanolic HCl and heated for 16 h at 80°C. The methanolysis product was dried, followed by the addition of 20 drops of methanol and drying 2 times. The samples were then, acetylated with pyridine and acetic anhydride in methanol in a 1:1:2 ratios at room temperature for 30 min, followed by per-O-trimethylsilylation with 200 μl of Tri-Sil (Pierce) and heating for 20 min. Then it was dried, dissolved in 2 ml of hexane, centrifuged, filtered through glass wool, and dried down to 100 μl. Aliquots of 1μl were analyzed by GC- MS (Agilent 6890N), using a HP-5m capillary column (Agilent 30 m x 0.25 mm) and mass selective detector (electron impact ionization mode). The oven temperature was programmed to increase from 80°C (2 min), to 140°C (2 min) at a rate of 20°C/min, to 200°C, at a rate of 2°C/min, and to 250°C, (5 min) at a rate of 30°C/min. The resulting peaks were identified by comparing their retention times with those of standard sugars using inositol as internal standard. For quantification, detector response factors (RF) were calculated for each standard sugar from peak area and weight of standard sugars and internal standard. RF values were used to calculate the weight of each component.

Molecular size was determined by gel filtration chromatography on Sephacryl S-400 column (1.6 cm x 90 cm) equilibrated and eluted with 0.1 M NaCl at a flow rate of 1ml/min. Polysaccharide fractions were monitored by phenol-sulfuric acid assay. Pullulan P-100 (100

KD), P-400 (376 KD), and P-800 (758 KD) were used as molecular size standards (Showa Denko K.K., Tokyo, Japan).

Glycosyl linkage analysis

The purified polysaccharides were methylated, hydrolyzed, reduced, and acetylated. The partially methylated alditol acetates (PMAAs) thus obtained were analyzed by GC-MS according to York et al. (1985). Aliquots of each polysaccharide (1.0 mg) were permethylated using the method of Ciucanu and Kerek (1984). The samples were suspended in 1 ml of DMSO, mixed with 0.7 ml of 1 M NaOH in DMSO, and incubated for 10 min. Then, 0.1 ml of methyl iodide was added to the suspension and it was incubated for 10 min. The permethylation was repeated twice using 0.2 ml of methyl iodide for 40 min to facilitate complete methylation of the polysaccharides in the second methylation. After that, the permethylated polysaccharides were extracted in organic phase of 1:1 dichloromethane-water, then hydrolyzed with 2 M trifluoroacetic acid (TFA) for 2 h at 121°C, and reduced with sodium borodeuteride (10 mg/ml in 1 M ammonia) overnight. The samples were neutralized with methanolic acetic acid, dried in methanol, and acetylated with 0.25 ml of acetic anhydride and 0.23 ml of TFA at 50°C for 10 min. After extraction into the organic phase of dichloromethane-Na_2CO_3, the PMAAs obtained were analyzed by GC-MS (Hewlett-Packard) using a Sp2330 capillary column (Supelco, 30 m x 0.25 mm) and mass selective detector (electron impact ionization mode). The oven temperature was programmed to increase from 80°C (2 min) to 170°C at a rate of 30°C/min and to 240°C (20 min) at a rate of 4°C/min. The resulted peaks were identified by comparing their retention times and mass spectra with those of PMAA of standard sugars, and quantified as mentioned above.

NMR spectroscopy

NMR analysis was done using a Bruker 400 M Hz spectrometer with a 5 mm sample tube, which was kept at 60°C. Polysaccharide samples were first exchanged with D_2O by dissolving in D_2O and lyophilizing 3 times to minimize the residual HDO. Standard pulse sequences from Bruker were used to obtain 1D proton, ^{13}C, 2D HSQC, COSY, TOCSY, and NOESY spectra. Chemical shifts were reported in ppm downfield from internal sodium 3-trimethylsilyl-(2,2,3,3,-$^{2}H_4$)-propanoate (TSP).

RESULTS

Glycosyl composition and molecular size analysis

Glycosyl compositional analysis using trimethylsilyl methylglycosides revealed that CPS of *A. tropicslis* SKU1100 (R) wild type strain was composed of glucose, galactose, and rhamnose in approximately equimolar ratio. Δ*polE* mutant EPS showed the same glycosyl composition as wild type (R) CPS. In gel filtration chromatography, EPS of Δ*polE* was eluted monodispersely earlier than that of wild type CPS (120 KD) (Moonmangmee et al., 2002b), and the molecular size of EPS was estimated to be around 620 KD.

Glycosyl linkage analysis

In order to know the sugar linkages of these polymers, methylation analysis of purified polysaccharides was carried out. As summarized in Table 2.1, wild type CPS contains 2,3-linked rhamnose residue, terminal and 2-linked glucose residues, terminal furanose residue, and 6-linked galactopyranose residue as major components, and some minor components, double branched 2,3,4-rhamnose, terminal galactopyranose, and 3-linked rhamnose (Fig. 2.1 A). EPS of Δ*polE* mutant had the same major components as those of the CPS and increased number of minor components such as glucose, galactose, and rhamnose residues (Fig. 2.1 B). The absence of the furanose residue was observed in Δ*galE* mutant EPS (see Chapter 3) suggesting that the terminal furanose residue observed in CPS and Δ*polE* EPS may be a galactofuranose as Δ*galE* mutant cannot produce galactose. To further check this possibility, we examined the electron impact ionization mass spectrum of the terminal furanose residue (Fig. 2.2 A) compared with the terminal galactopyranose residue (Fig. 2.2 B). The mass spectrum of the former peak was different from that of the latter peak, but rather the same Spectrum as terminal galactofuranose electron impact ionization spectrum (Sassaki et al., 2005). Thus, these results confirmed that the terminal furanose residue was a galactofuranose.

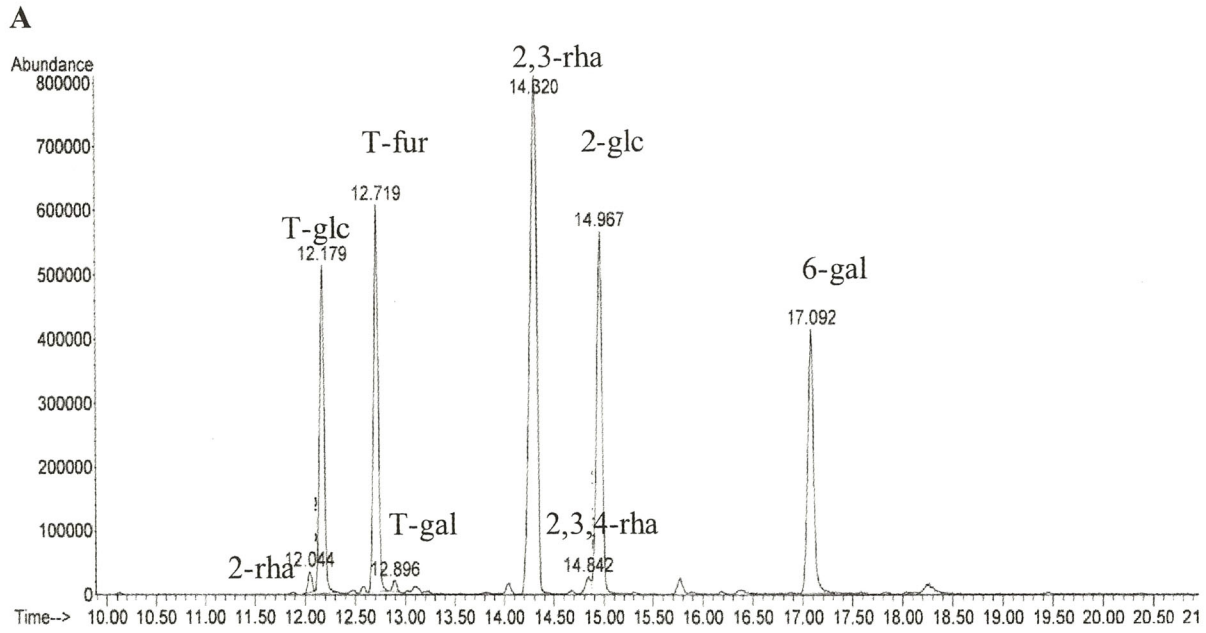
NMR analysis

Proton and ^{13}C NMR spectroscopies were used to determine the number of residues in the repeating units in *A. tropicalis* SKU1100 polysaccharides. The 1D proton spectrum of *A. tropicalis* SKU1100 wild type CPS showed five signals (A/B, C, D, E, and G in Fig. 2.3 A; A/B are overlapped as shown in Fig. 2.4) in the anomeric region (4.4-5.5 ppm), crowded ring region, and one signal corresponding to methyl group of 6-deoxysugars (CH_3 in Fig. 2.3 A).

Table 2.1 Methylation analysis of *A. tropicalis* SKU1100 wild type CPS, and Δ*polE* mutant EPS.

	Moles%	
Sugar residue	**SKU1100 (R)**	**Δ*polE* mutant**
2,3-rhamnose	33.5	28.5
2-glucose	18.6	18.1
Terminal- galactofuranose	17.3	18.1
6-galactose	14.1	12.8
Terminal-glucose	14.8	17.0
2,3,4-rhamnose	0.69	NF
3-rhamnose	0.34	1.70
Terminal-galactose	0.34	3.10
6-glucose	NF	0.34
3-glucose	NF	0.34
3,4-glucose	NF	0.34
2,6-glucose	NF	0.34
2-rhamnose	NF	0.20
4,6-galactose	NF	0.10
4-glucose	NF	NF
2,3-glucose	NF	NF

[a]NF: not foun.



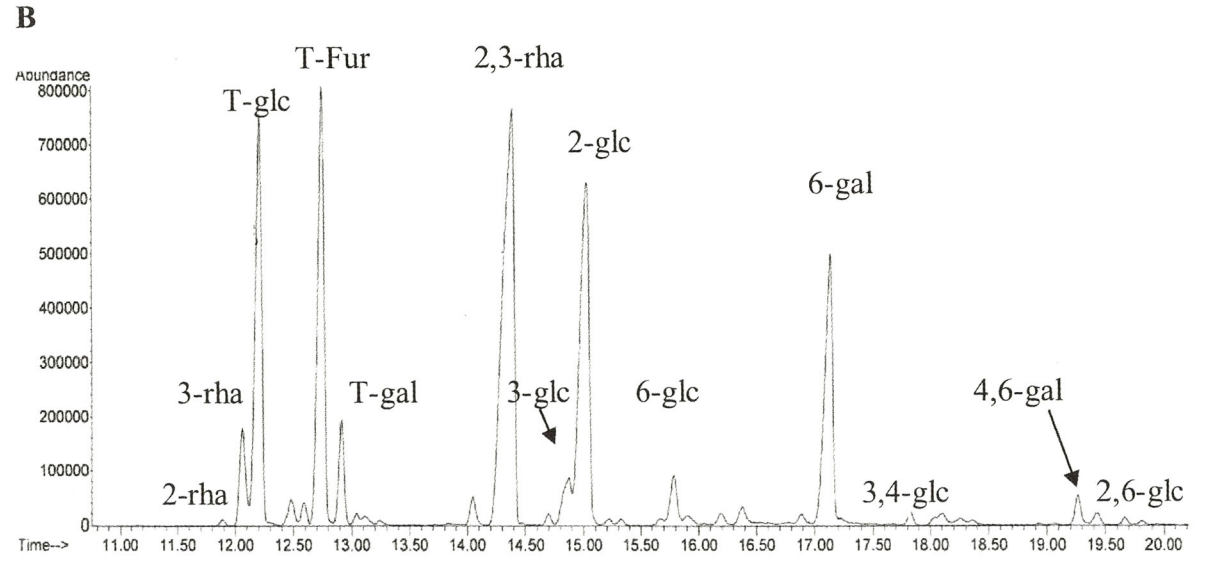
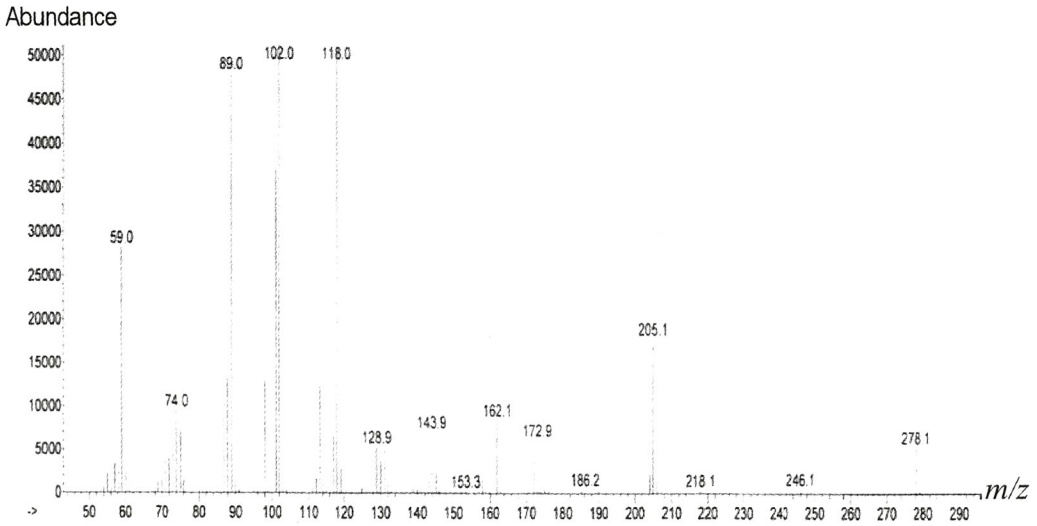


Fig. 2.1 GC-MS analysis of partially methylated alditol acetates of *A. tropicalis* SKU1100 wild type CPS (A), and Δ*polE* mutant EPS (B).

T: terminal, glc: glucose, gal: galactose, Fur: hexofuranose, and rha: rhamnose.

A



B

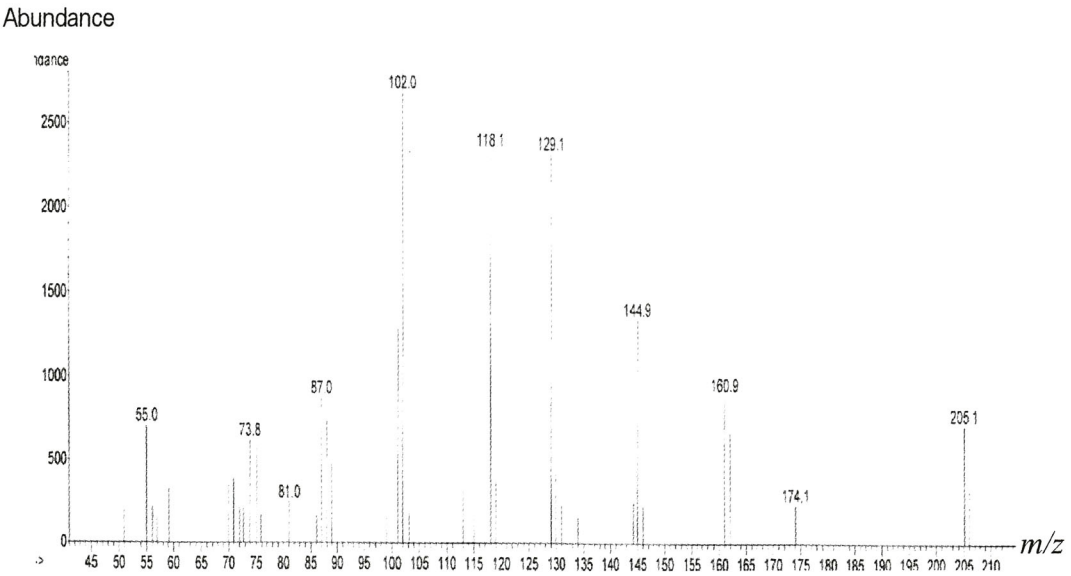


Fig. 2.2 Electron impact ionization mass spectra of terminal galactofuranose residue (A) and of terminal galactopyranose residue (B) of *A. tropicalis* SKU1100 (R) wild-type CPS.

The analysis was carried out with the two peaks of retention times, 12.719 min of a terminal furanose residue (T-Fur) and 12.896 min of a terminal galactose residue (T-Gal), in GC-MS analysis of partially methylated alditol acetates of *A.tropicalis* SKU1100 wild-type CPS (Fig. 2.1 A). As described in the text, the two peaks (T-Fur and T-Gal) are regarded as Terminal-galactofuranose and Terminal-galactopyranose, respectively (Table 2.1).

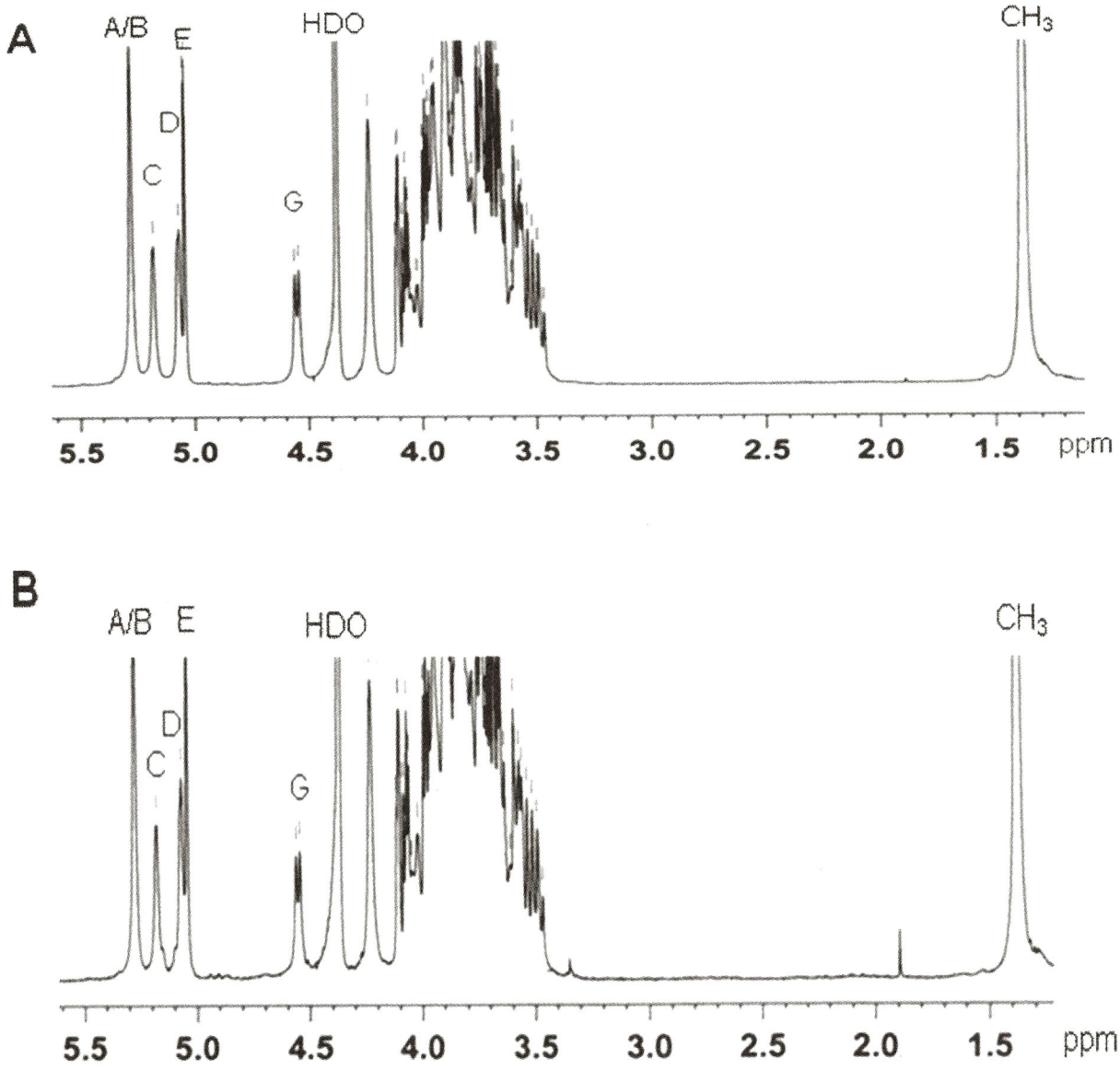


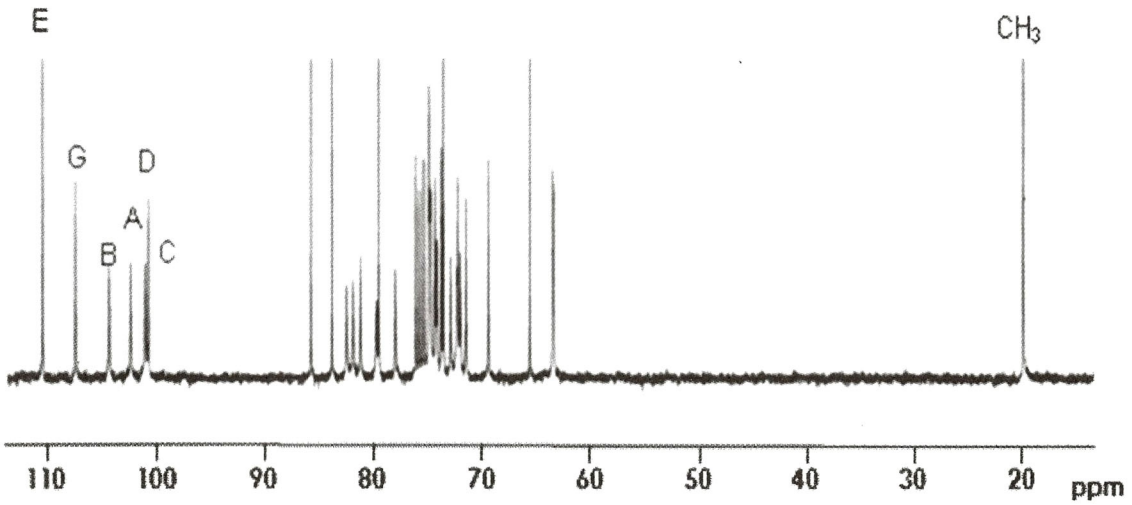
Fig. 2.3 Proton NMR spectra of *A. tropicalis* SKU1100 wild-type (R) CPS (A), and Δ*polE* mutant EPS (B).

The spectra were recorded in D_2O at 60°C with TSP as internal standard. The peaks named as A to G as described in the text.

EPS of Δ*polE* mutant exhibited a typical 1D proton spectrum as much the same as wild type CPS (Fig. 2.3 B). Furthermore, ^{13}C NMR analysis exhibited six anomeric signals (A, B, C, D, E, and G) in both wild type CPS and Δ*polE* mutant EPS (Fig. 2.4 A and B). Since anomeric signal A in ^{1}H spectrum showed correlation to two signals of ^{13}C in HSQC spectrum (A and B, data not shown), A and B signals were shown to be overlapped in ^{1}H spectrum (Fig. 2.3 A

and B). Together with the methylation analysis results, these findings implied that SKU1100 wild type CPS and Δ*polE* mutant EPS had the same hexasaccharide repeating unit.

A



B

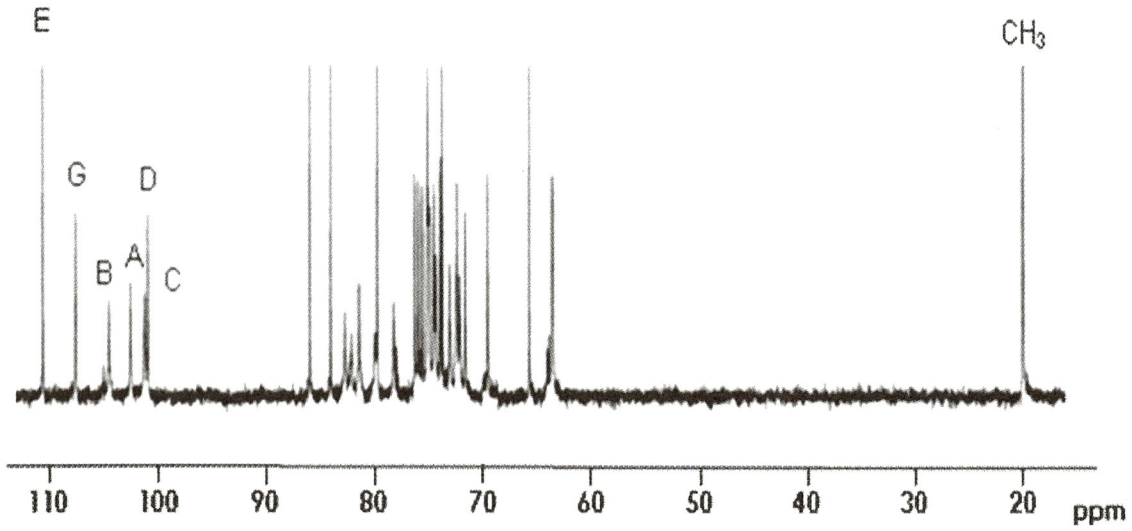


Fig. 2.4 ^{13}C NMR spectra of *A. tropicalis* SKU1100 wild-type (R) CPS (A), and Δ*polE* mutant EPS (B).

The peaks named as in Fig. 2.3.

The 2D ^{1}H-^{1}H COSY, TOCSY, and ^{1}H,^{13}C HSQC experiments, alongside with 2D NOESY, 1D ^{1}H, and ^{13}C analysis, were used to assign the proton and carbon chemical shifts, absolute configurations, and anomeric configurations of the polysaccharides. Proton chemical shifts assignment was hindered by signal overlapping in the ring region of SKU1100 wild type CPS and Δ*polE* EPS. However, together with the analysis of ^{13}C-NMR and ^{1}H-^{13}C HSQC, the analysis of Δ*galE* mutant which has fewer number of residues than wild type CPS and Δ*polE* mutant EPS (see Chapter 3) enabled us to complete the signal assignment. Anomeric proton and carbon signals labeled as A to E and G in wild type CPS (Fig. 2.3 and 2.4 A), and Δ*polE* mutant EPS (Fig. 2.3 and 2.4 B), from downfield upwards were identified by comparing their chemical shifts (ppm) with those reported in the literature (Roslund et al., 2008; Hounsell, 1995; Perepelov et al., 2010; Landersjö et al., 2002; Grimmecke et al., 1991; Gargiulo et al., 2008). As shown in Table 2.2, SKU1100 wild type CPS constituents were identified as 2-linked-α-D-glucopyranose, terminal-α-D-glucopyranose, 2,3-linked-α-L-rhamnopyranose (2 residues), terminal-β-D-galactofuranose, and 6-linked-β-D-galactopyranose residues. EPS of Δ*polE* mutant showed the same sugar residues and configuration as wild type CPS.

This identification was confirmed by ^{13}C assignment (Table 2.3) which showed the same constituents as mentioned above. Moreover, the characteristic down field shift of glycosylated carbon peaks confirmed the linkage analysis results. The small $J_{H1, H2}$ coupling constants (J=2.2, 2.8 Hz) suggested that A and D residues were in α configuration. The anomeric carbon E resonated at 110.41 ppm (Table 2.3) indicating furanosidic ring form, and the small coupling (J=1.6) confirmed its β-furanose configuration. Residue G showed large coupling constant (J=7.5) of β-pyranose configuration (Table 2.2).

Figure 2.5 shows the inter-residual correlations between the anomeric protons and nearby protons giving information about saccharides sequence in repeating unit. SKU1100 wild type CPS exhibited cross peaks between anomeric proton G and H-2 of A, indicating the link between A and G, and further did that anomeric protons of C and E show a correlation to H-2 of B and D, indicating that C and E are branches. For Δ*polE* mutant EPS, similar cross peaks were observed (Fig 2.5 A and B).

These correlations, together with methylation analysis results and comparison between wild type CPS and Δ*galE* mutant EPS (see chapter 3), allowed us to estimate the

monosaccharide sequences of the polysaccharides (the minimum repeating units), as shown in Fig. 2.6.

Table 2.2 Proton chemical shifts of *A.tropicalis* SKU1100 CPS, and Δ*polE* mutant EPS. $J_{H1,H2}$ coupling constants (Hz) included in braces.

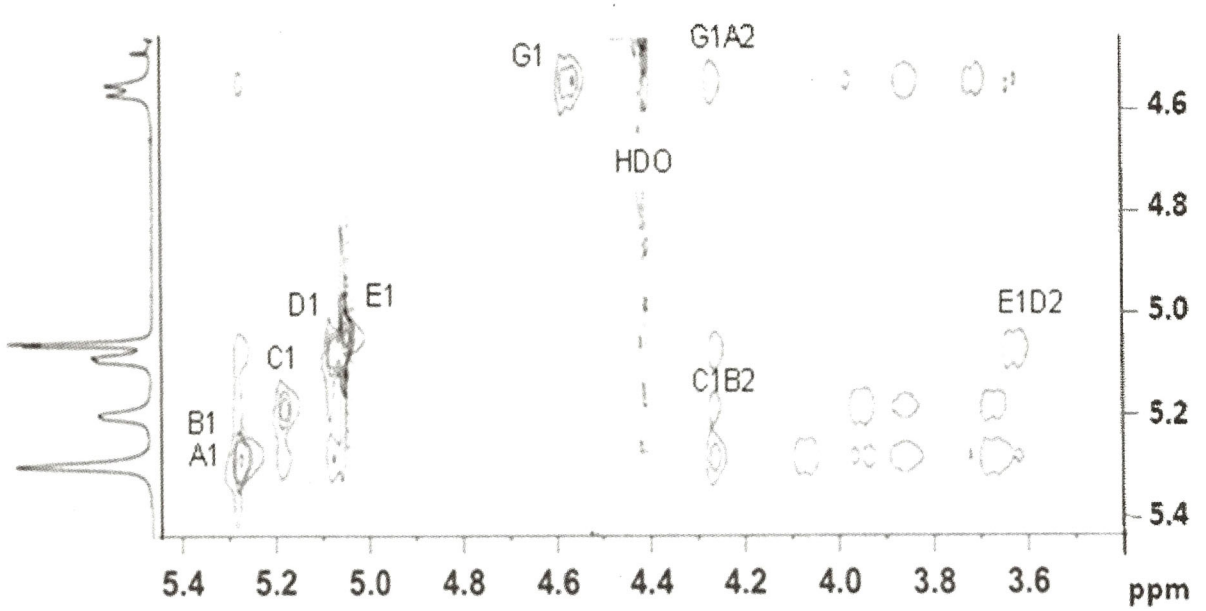
	Chemical shifts(ppm)					
Sugar residue	**H-1**	**H-2**	**H-3**	**H-4**	**H-5**	**H-6**
CPS from SKU1100 wild type						
A→ 2)-α-D-Glcp-(1→	5.28{2.2}	4.23	4.15	3.82	3.63	3.7/3.72
B→2,3)-α-L-Rhap-(1→	5.28	4.23	4.02	3.60	3.8	1.38
C –α-D-Glcp-(1→	5.18	3.61	3.91	3.44	3.57	3.79/3.77
D → 2,3)-α-L-Rhap-(1→	5.07{2.8}	3.65	3.97	3.61	3.93	1.37
E -β-D-Galf-(1→	5.04{1.6}	4.11	4.08	4.06	3.98	3.73
G→ 6)-β-D-Galp-(1→	4.56{7.5}	3.71	3.63	3.98	3.81	3.59
EPS from Δ*polE* mutant						
A → 2)-α-D-Glcp-(1→	5.28{2.2}	4.23	4.15	3.82	3.61	3.65/3.72
B→2,3)-α-L-Rha-(1→	5.28	4.23	4.02	3.61	3.8	1.38
C –α-D-Glcp-(1→	5.18	3.60	3.94	3.47	3.57	3.79/3.75
D → 2,3)-α-L-Rhap-(1→	5.07{2.8}	3.65	3.97	3.63	3.95	1.38
E -β-D-Galf-(1→	5.04{1.6}	4.11	4.06	3.98	3.73	3.56
G→ 6)-β-D-Galp-(1→	4.56{7.5}	3.71	3.63	3.98	3.81	3.59

[a] chemical shift values were rounded to two digits.

In wild type CPS structure, there was an additional 2,3,4-α-L-rhamnopyranose double branched residue albeit with small amount (Table 2.1), which could not be seen in ΔpolE mutant EPS. Thus, as discussed later, it could be possible that this residue occurs in only part of many repeating units, together with another minor residues such as β-D-galactopyranose

branch residue (in box in Fig. 2.6 A), which could not be detected in NMR analysis because of its low abundancy (Table. 2.1).

A



B

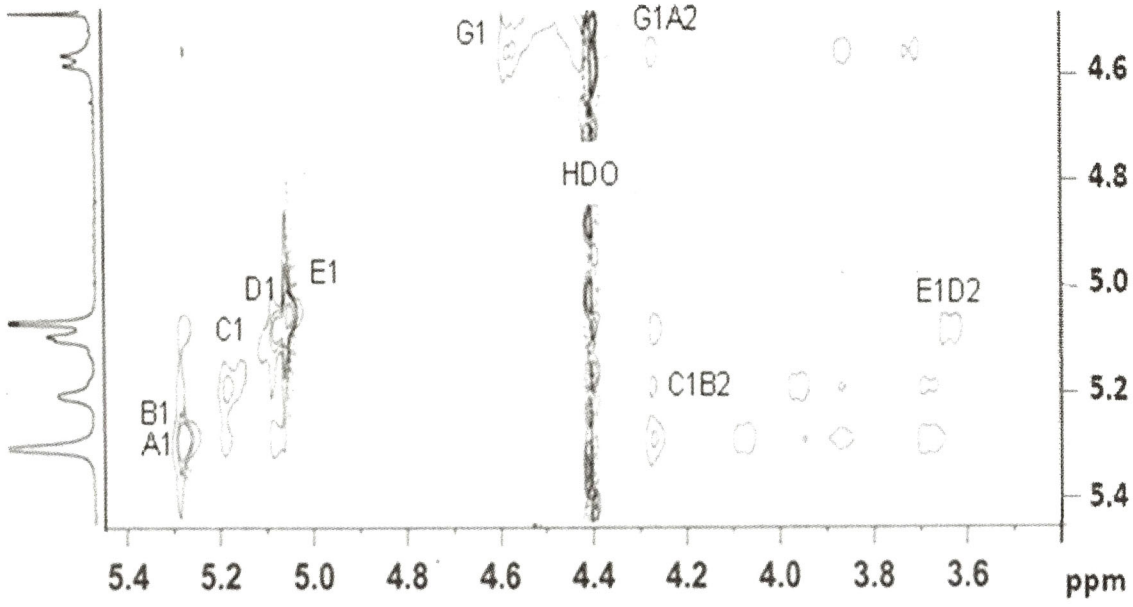


Fig. 2.5 Part of NOESY spectra of *A. tropicalis* SKU1100 wild-type (R) CPS (A), and Δ*polE* mutant EPS (B).

Table 2.3 ^{13}C chemical shifts of *A.tropicalis* SKU1100 CPS, and Δ*polE* mutant EPS.

	Chemical shifts (ppm)					
Sugar residue	**C-1**	**C-2**	**C-3**	**C-4**	**C-5**	**C-6**
CPS from SKU1100 wild type						
A→ 2)-α-D-Glcp-(1→	102.35	78.15	74.47	75.89	74.89	63.46
B→2,3)-α-L-Rhap-(1→	104.31	81.35	82.61	72.08	74.31	19.91
C-α-D-Glcp-(1→	101.07	72.33	71.52	73.02	74.89	63.52
D → 2,3)-α-L-Rhap-(1→	100.76	83.94	79.89	75.51	74.31	19.88
E -β-D-Galf-(1→	110.41	82.03	85.86	72.42	76.25	65.63
G→ 6)-β-D-Galp-(1→	107.41	75.01	79.67	73.70	73.88	69.46
EPS from Δ*polE* mutant						
A→ 2)-α-D-Glcp-(1→	102.36	78.16	74.47	75.90	74.89	63.46
B→2,3)-α-L-Rhp(1→	104.32	81.37	82.62	72.08	74.32	19.92
C -α-D-Glcp-(1→	101.08	72.33	71.53	73.03	74.89	63.52
D→2,3)-α-L-Rhap-(1→	100.77	83.95	79.90	75.52	74.32	19.88
E -β-D-Galf-(1→	110.42	82.03	85.87	72.34	76.26	65.64
G → 6)-β-D-Galp-(1→	107.42	75.02	79.68	73.71	73.89	69.47

DISCUSSION

We identified the structure of *A. trpoicalis* SKU1100 wild type CPS, as well as Δ*polE* mutant EPS. Wild type CPS and Δ*polE* mutant EPS appears to have the same hexasaccharide repeating unit. Monosaccharide composition analysis of wild type CPS, and Δ*polE* EPS show similar results as the previous studies (Deeraksa, et al., 2005; Moonmangmee, et al., 2002b). The results obtained here showed no acidic residue in these CPS and EPS, by using different

method (TMS) which could detect acidic residues if any, indicating the neutral nature of these polysaccharides.

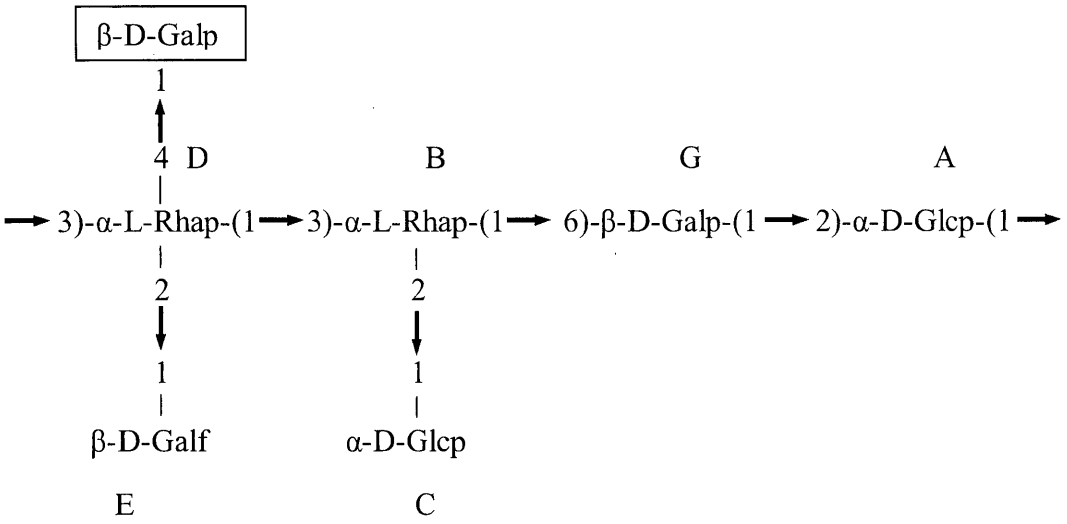
Methylation linkage analysis showed increased number of minor glucose, galactose, and rhamnose residues in Δ*polE* mutant EPS. These minor residues, detected only by GC-MS analysis, may be present in the edges of the CPS or EPS polymer complex. Especially, one minor residue 2,3,4-α-L-rhamnose, detected only in wild type CPS but not in the mutant EPS, would be important to consider the difference in the localization between CPS and EPS. Since this double branched residue represents only about 0.69% of the CPS composition (Table 2.1), it may occurs only once or few times in large numbers of the repeating units of the CPS structure. Thus, it can be speculated that this double branched residue may play a role in CPS formation, in which the modification of the 2,3-α-L-rhamnose inside the repeating units may terminate the polymerization reaction, and lead the polysaccharide binding to the cell surface. Since actually Δ*polE* mutant EPS has much larger size than wild type CPS, the continuous polymerization reaction would occur to make much larger polymer in the absence of PolE which may cause such the termination reaction with the minor residues.

It is now well established that the polysaccharide co-polymerase proteins (PCP) regulate the chain length and surface assembly of capsular polysaccharides (Collins, et al., 2007). Although the PCP homologue could be found in the genome of *Gluconacetobacter diazotrophicus* (GDI_2493) but not of *Acetobacter pasteurianus,* however, *polE* gene showed no similarity to PCP protein family. Whereas, such the homology search of *polE* gene showed that there are homologous *polE* genes in many bacteria including acetic acid bacteria, e.g. putative rhamnosyltransferases from *Acetobacter pasteurianus, Granulobacter bethesdensis,* and putative glycosyltransferases from *Gluconobacter oxydans* 621H, *Rhodobacter sphaeroides*, and *Nitrosomonas eutropha* C71 (data not shown). Moreover, PolE has clear consensus motifs (catalytic domain), DDGSxD and DQDDxW, near the N-terminus, and additional consensus motifs (estimated as the binding site), HDWxx and xYRQH, near the central part. The former two motifs are similar to domain A of ExoU and HasA families of β-glycosyltransferases, but the latter two are different from the putative substrate binding domain of this family (Deeraksa, et al., 2005).

Thus, taking together with the structural findings, we propose that PolE may work as β-galactosyltransferase that adds terminal-β-galactopyranosyl residue to 2,3-α-L-

rhamnopyranosyl residue to make 2,3,4-α-rhamnopyranosyl residue (Fig. 2.6 A). This notion could be supported by the following observation. Like Δ*polE* mutant, despite of keeping native *polE* gene, Δ*galE* mutant also produces EPS but not CPS (see chapter 3). Since the mutant could not produce UDP-galactose as the precursor of the terminal-β-galactopyranosyl residue, active PolE could not make the double branched rhamnosyl residue and thus not able to switch the EPS repeating units to the CPS. Thus, it is conceivable that this double branched residue contribute to CPS attachment to the cell surface. However, biochemical characterization of PolE remains to be addressed.

A



B

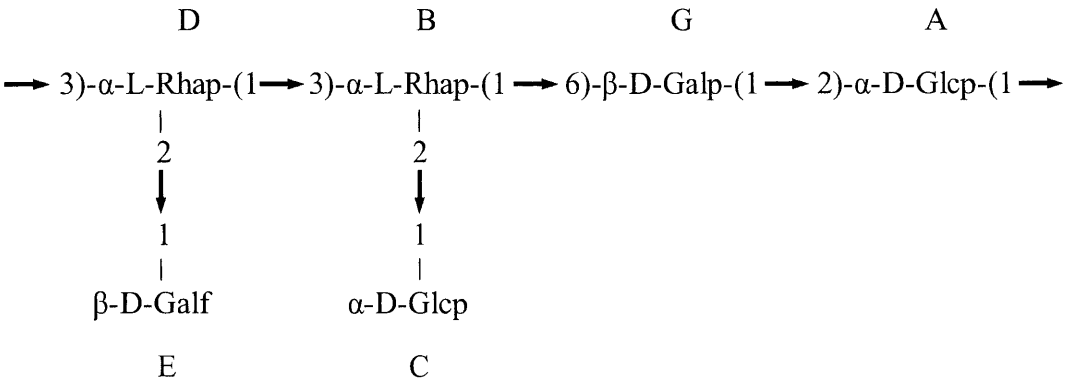


Fig. 2.6 Proposed repeating unit structures of *A. tropicalis* SKU1100 wild-type (R) CPS (A), and Δ*polE* mutant EPS (B).

A to G on each residue in the repeating units refer to the residues named in Fig. 2.3. The β-D-galactopyranosyl residue in box is not a part of the repeating unit of wild type CPS, it rather

substitutes 2,3-L-rhamnosyl residue to make 2,3,4-L-rahamnosyl double branch which occurred only once or few times in the whole structure of the CPS.

Since the conversion of CPS to EPS occurs spontaneously in the case of S strain besides the artificial mutations described, this indicates the ability of *A. tropicalis* to produce different types of polysaccharides in different niches. For example, in shaking submerge culture, membrane stresses would be generated more than in static culture by the higher acid production due to rapid growth, and thus EPS production may help in stress adaptation (Mao et al., 2001). Recently, CPS, and probably EPS, production has been shown to add acetic acid resistance to *A. tropicalis* (Deeraksa et al., 2005) and *A. pasteurianus* (Kanchanarach et al., 2010). Additionally, *A. tropicalis* was recently reported as a major symbiont in the olive fruit fly (Kounatidis et al., 2009), reasonably suggesting that A. *tropicalis* produce EPS when switching from free-living mode to symbiosis mode to aid its colonization of the host insect.

In conclusion, we have determined the structure of *A. tropicalis* CPS, based on the structural comparison with mutant EPS, to be a polymer consisting of a branched hexasaccharide repeating unit composed of two moles 2,3-α-L-Rhap, and one mole each of 6-β-D-Galp and 2-α-D-Glcp, of which the two rhamnosyl residues are branched by β-D-Galf and α-D-Glcp. Furthermore, the study could confer information physiologically important for *Acetobacter* sp., which is directly related to the pellicle formation. EPS could be attached to cell surface by the function of PolE, which may cross-link one 2,3-α-L-Rhap with β-D-Galp to produce CPS.

CHAPTER 3

Structure of exopolysaccharide produced by *Aetobacter tropicalis* SKU1100 Δ*galE* mutant

ABSTRACT

Acetobacter tropicalis SKU1100 (R) strain produces capsular polysaccharide (CPS) that involves in pellicle formation and growth in static culture. Compositional analysis of the CPS showed that it is composed of galactose, glucose, and rhamnose. Moreover, *galE* gene (UDP-galactose 4-epimerase) was found to be essential for CPS production and pellicle formation in *A. tropicalis* SKU1100, since disruption of this gene abolishes the pellicle formation and CPS production. The Δ*galE* mutant produced EPS with higher molecular weight than CPS that secreted in the growth medium. In this study, structural characterization of the EPS was done using monosaccharide composition, methylation analysis, and 1D and 2D NMR. The EPS contained glucose and rhamnose in approximately equimolar ratio and had a branched tetrasaccharide repeating unit composed of 2,3-α-L-rhamnopyranose, 3-α-L-rhamnopyranose, 2-α-D-glucopyranose, and terminal-α-D-glucopyranose branch residues. The absence of galactose residues (terminal-β-D-galactofuranose and 6-β-D-galactopyranose) that are identified in wild type CPS was observed, confirming the function of *galE* gene in producing UDP-galactose from UDP-glucose.

INTRODUCTION

Many bacteria can produce exopolysaccharides (EPSs) that have high molecular weight, secreted in to the growth medium, and are not attached to the cell surface. Production of slime or mucoid colony morphology is directly associated with EPS production in many microorganisms (Laws et al., 2001). EPSs protect microbes in their natural environments from desiccation and stresses (Ferreira et al., 2010), while their application in food industry add thickening and gelling properties (Sutherland, 1998). EPSs are either homopolymers composed of one type of monosaccharides or heteropolymers that contain more than one type of monosaccharides. The functionality of the EPS is determined by its primary structure which specifies the folding type and association under various conditions.

Galactose metabolism in Gram-negative bacteria is catalyzed by the Leloir pathway enzymes galKTE (Metzeger et al., 1994) of which *galE* gene encoded, UDP-galactose 4-epimerase enzyme, converts UPD-galactose to UDP-glucose and vice versa. *galE* gene is involved in lipopolysaccharides biosynthesis and virulence in many bacteria (Fry et al., 2000; Jennings et al., 1993). Moreover, *galE* overexpression in thermophilic bacteria *Thermus thermophilus* led to increase in biofilm formation because of its involvement in exopolysaccharide biosynthesis (Niou et al., 2009). *Gulconacetobacter xylinus* produces EPS called acetan in addition to cellulose production. Acetan has similar structure as commercially important polysaccharide xanthan thus, a new exopolysaccharide (P2), was produced by deleting a glycosyltransferase gene (*aceP)* from acetan producing *A. xylinum* strain CKE5 (Colquhoun et al., 2001). *A. tropicalis* SKU1100 R strain is capable of producing pellicle when grown statically, capsular polysaccharide production is essential for pellicle formation. In previous study, Deeraksa and co workers (2005) identified gene operon *polABCDE* important for CPS biosynthesis and pellicle formation in *A. tropicalis* SKU1100. *polABCD* showed high similarity to *rfbBACD* which is responsible for d-TDP-rhamnose biosynthesis while *polE* is a novel gene with relatively low similarity to glycosyltransferases. In addition, *galE*, a gene that encodes UDP-galactose 4-epimerase involved in UDP-galactose biosynthesis, has also been identified in *A. tropicalis* SKU1100. The disruption of this gene resulted in no pellicle formation in static culture but secretion of EPS composed of glucose and rhamnose in culture medium. This EPS was found to have higher molecular weight and viscosity than wild type CPS (Deeraksa et al., 2006).

In this investigation we set out to elucidate the structure of Δ*galE* mutant strain EPS and thereby gain insight into *galE* function.

MATERIALS and METHODS

Bacterial strains and culture conditions

Δ*galE* mutant of *A. tropicalis* SKU1100 (R) strain (Deeraksa et al., 2006) was used in this study. The *galE* gene was disrupted by insertion of non-polar kanamycin resistant cassette. The strain was grown at 30°C in potato medium (Deeraksa et al., 2005). Kanamycin was added at the concentration of 50 μg/ml.

Purification of Δ*galE* mutant EPS

The EPS of Δ*galE* mutant was purified from culture media by essentially the same method as described (Deeraksa et al., 2006). The cultivation was performed in 1 L of YPG medium (5% glycerol, 0.5% peptone, and 0.5% yeast extract) for 2 days. The cells were removed by centrifugation (9000 x *g* for 10 min at 4°C). The culture media was then collected and concentrated to one third using ultra filtration (20 KD cutoff, Advantec), followed by DEAE-cellulose column chromatography as described (Chapter 2). Polysaccharide fractions were then pooled, and treated as described (Chapter 2), and applied to a Sephacryl S-400 column equilibrated with 0.1 M NaCl. Polysaccharide fractions were combined and precipitated with 2 volumes of cold ethanol.

Monosaccharide composition and molecular size analysis

The monosaccharide composition was analyzed using trimethylsilyl (TMS) methylglycoside method as described by Wozniak et al. (2003). Aliquot of 300-μg of the polysaccharide was lyophilized, and then, mixed with 500 μl of 1M methanolic HCl and heated for 16 h at 80°C. The methanolysis product was dried, followed by the addition of 20 drops of methanol and drying 2 times. The sample was then, acetylated with pyridine and acetic anhydride in methanol in a 1:1:2 ratios at room temperature for 30 min, followed by per-O-trimethylsilylation with 200 μl of Tri-Sil (Pierce) and heating for 20 min. Then it was dried, dissolved in 2 ml of hexane, centrifuged, filtered through glass wool, and dried down to 100 μl. Aliquots of 1μl were analyzed by GC- MS (Agilent 6890N), using a HP-5m capillary column (Agilent 30 m x 0.25 mm) and mass selective detector (electron impact

ionization mode). The oven temperature was programmed to increase from 80°C (2 min), to 140°C (2 min) at a rate of 20°C/min, to 200°C, at a rate of 2°C/min, and to 250°C, (5 min) at a rate of 30°C/min. The resulting peaks were identified by comparing their retention times with those of standard sugars using inositol as internal standard. For quantification, detector response factors (RF) were calculated for each standard sugar from peak area and weight of standard sugars and internal standard. RF values were used to calculate the weight of each component.

Molecular size was determined by gel filtration chromatography on Sephacryl S-400 column (1.6 cm x 90 cm) equilibrated and eluted with 0.1 M NaCl at a flow rate of 1ml/min. Polysaccharide fractions were monitored by phenol-sulfuric acid assay. Pullulan P-100 (100 KD), P-400 (376 KD), and P-800 (758 KD) were used as molecular size standards (Showa Denko K.K., Tokyo, Japan).

Glycosyl linkage analysis

The purified EPS was methylated, hydrolyzed, reduced, and acetylated. The partially methylated alditol acetates (PMAAs) thus obtained were analyzed by GC-MS according to York et al. (1985). Polysaccharide (1.0 mg) was permethylated using the method of Ciucanu and Kerek (1984). The sample was suspended in 1 ml of DMSO, mixed with 0.7 ml of 1 M NaOH in DMSO, and incubated for 10 min. Then, 0.1 ml of methyl iodide was added to the suspension and it was incubated for 10 min. The permethylation was repeated twice using 0.2 ml of methyl iodide for 40 min to facilitate complete methylation of the polysaccharide in the second methylation. After that, the permethylated polysaccharide was extracted in organic phase of 1:1 dichloromethane-water, then hydrolyzed with 2 M trifluoroacetic acid (TFA) for 2 h at 121°C, and reduced with sodium borodeuteride (10 mg/ml in 1 M ammonia) overnight. The sample was neutralized with methanolic acetic acid, dried in methanol, and acetylated with 0.25 ml of acetic anhydride and 0.23 ml of TFA at 50°C for 10 min. After extraction into the organic phase of dichloromethane-Na_2CO_3, the PMAAs obtained were analyzed by GC-MS (Hewlett-Packard) using a Sp2330 capillary column (Supelco, 30 m x 0.25 mm) and mass selective detector (electron impact ionization mode). The oven temperature was programmed to increase from 80°C (2 min) to 170°C at a rate of 30°C/min and to 240°C (20 min) at a rate of 4°C/min. The resulted peaks were identified by comparing their retention times and mass spectra with those of PMAA of standard sugars, and quantified as mentioned above.

NMR spectroscopy

NMR analysis was done using a Bruker 400 M Hz spectrometer with a 5 mm sample tube, which was kept at 60°C. EPS was first exchanged with D_2O by dissolving in D_2O and lyophilizing 3 times to minimize the residual HDO. Standard pulse sequences from Bruker were used to obtain 1D proton, ^{13}C, 2D HSQC, COSY, TOCSY, and NOESY spectra. Chemical shifts were reported in ppm downfield from internal sodium 3-trimethylsilyl-(2,2,3,3,-$^2H_4$)-propanoate (TSP).

RESULTS

Glycosyl composition and molecular size analysis

Glycosyl compositional analysis using trimethylsilyl methylglycosides revealed that EPS of Δ*galE* mutant was composed of glucose and rhamnose in approximately equimolar ratio. In gel filtration chromatography, EPS of Δ*galE* mutant was eluted monodispersely at almost the same elution volume as Δ*polE* mutant EPS, which is earlier than that of wild type CPS (120 KD) (Moonmangmee et al., 2002b), and the molecular size of the EPS was estimated to be the same as Δ*polE* mutant EPS (620 KD).

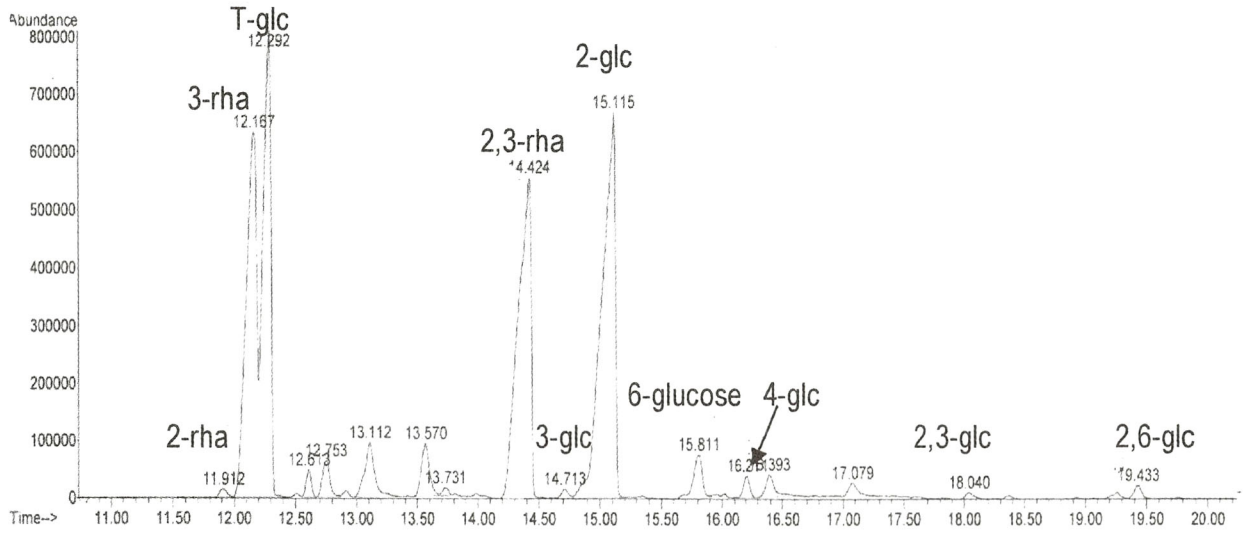


Fig. 3.1 GC-MS analysis of partially methylated alditol acetates of Δ*galE* mutant EPS.

T: terminal, glc: glucose, and rha: rhamnose.

Glycosyl linkage analysis

In order to know the sugar linkages of this polymer, methylation analysis of purified EPS was carried out. As summarized in Table 3.1, Δ*galE* mutant EPS contained 2,3-linked rhamnose, 3-linked rhamnose, terminal and 2-linked glucose residues as the major components and some minor components of glucose, and rhamnose (Fig. 3.1). The absence of galactose residues (terminal galatofuranose and 6-galactopyranose in wild type CPS and Δ*polE* mutant EPS) was observed in Δ*galE* mutant EPS and this is in consistent with the disruption of *galE* gene, as the function of *galE* gene is to produce UDP-galactose from UDP-glucose.

NMR analysis

Proton and ^{13}C NMR spectroscopies were used to determine the number of residues in the repeating units in Δ*galE* polysaccharide. The 1D proton spectrum of Δ*galE* mutant EPS showed four signals (A, C, D, and F in Fig. 3.2) in the anomeric region, and the characteristic methyl group signal of 6-deoxysugars. Furthermore, ^{13}C NMR analysis of Δ*galE* mutant EPS showed the same four anomeric signals (A, C, D, and F in Fig.3.3) as in ^{1}H spectrum (Fig. 3.2). Together with the methylation analysis results, these findings suggested that, Δ*galE* mutant EPS had a tetrasaccharide repeating unit.

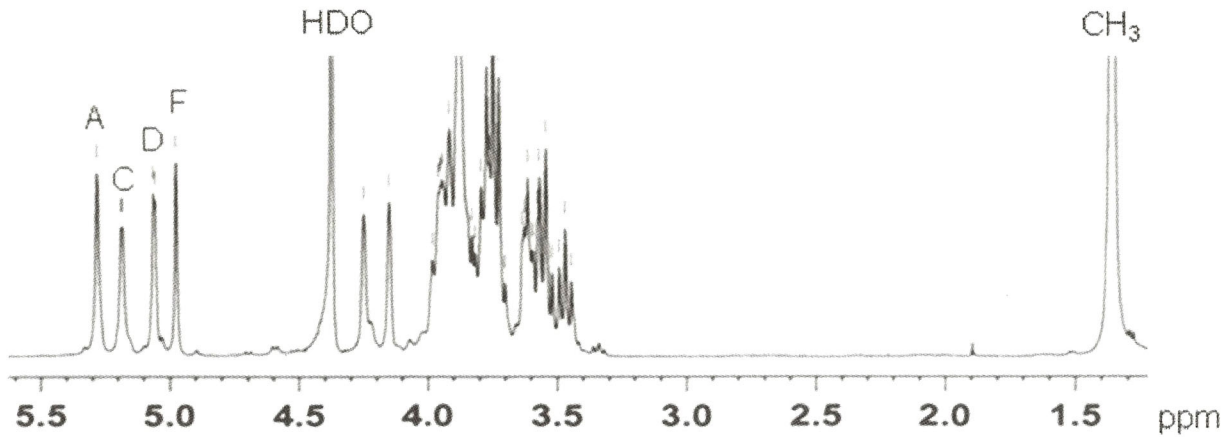


Fig. 3.2 Proton NMR spectrum of Δ*galE* mutant EPS.

The spectrum was recorded in D_2O at 60°C with TSP as internal standard. The peaks named as described in the text.

The 2D ^{1}H-^{1}H COSY, TOCSY, and ^{1}H,^{13}C HSQC experiments, alongside with 2D NOESY, 1D ^{1}H, and ^{13}C analysis, were used to assign the proton and carbon chemical shifts, absolute configurations, and anomeric configurations of the polysaccharides. Anomeric proton and carbon signals labeled as A, C, D, and F in Δ*galE* mutant EPS (Fig. 3.2 and 3.3) from downfield upwards were identified by comparing their chemical shifts (ppm) with those reported in the literature (Roslund et al., 2008; Hounsell, 1995; Perepelov et al., 2010; Landersjö et al., 2002; Grimmecke et al., 1991; Gargiulo et al., 2008).

Table 3.1 Methylation analysis of Δ*galE* mutant EPS.

	Moles%
Sugar Residue	**Δ*galE* mutant**
2,3-rhamnose	23.9
2-glucose	28.0
Terminal- galactofuranose	NF
6-galactose	NF
Terminal-glucose	21.5
2,3,4-rhamnose	NF
3-rhamnose	24.2
Terminal-galactose	NF
6-glucose	0.34
3-glucose	0.30
3,4-glucose	NF
2,6-glucose	0.34
2-rhamnose	0.34
4,6-galactose	NF
4-glucose	0.68
2,3-glucose	0.23

[a]NF: not found

As shown in Table 3.2, Δ*galE* mutant EPS contained 2-linked-α-D-glucopyranose, terminal-α-D-glucopyranose, 2,3-linked-α-L-rhamnopyranose, and 3-linked-α-L-rhamnopyranose. This identification was confirmed by ^{13}C assignment (Table 3.3) which showed the same constituents as mentioned above. Moreover, the characteristic down field shift of glycosylated carbon peaks confirmed the linkage analysis results. The small $J_{H1, H2}$ coupling constants (J=2.2, 2.8 Hz) suggested that A and D residues were in α configuration.

Table 3.2 Proton chemical shifts of Δ*galE* mutant EPS.

$J_{H1,H2}$ coupling constants (Hz) included in braces.

	Chemical shifts(ppm)					
Sugar residue	**H-1**	**H-2**	**H-3**	**H-4**	**H-5**	**H-6**
A → 2)-α-D-Glcp-(1→	5.28{2.2}	4.25	3.91	3.59	3.63	3.88/3.76
C –α-D-Glcp-(1→	5.19	3.61	3.91	3.44	3.57	3.79/3.77
D → 2,3)-α-L-Rhap-(1→	5.06{2.8}	3.72	3.62	3.96	3.94	1.36
F → 3)-α-L-Rhap-(1→	4.98	4.15	3.95	3.57	3.7	1.35

[a] chemical shift values were rounded to two digits.

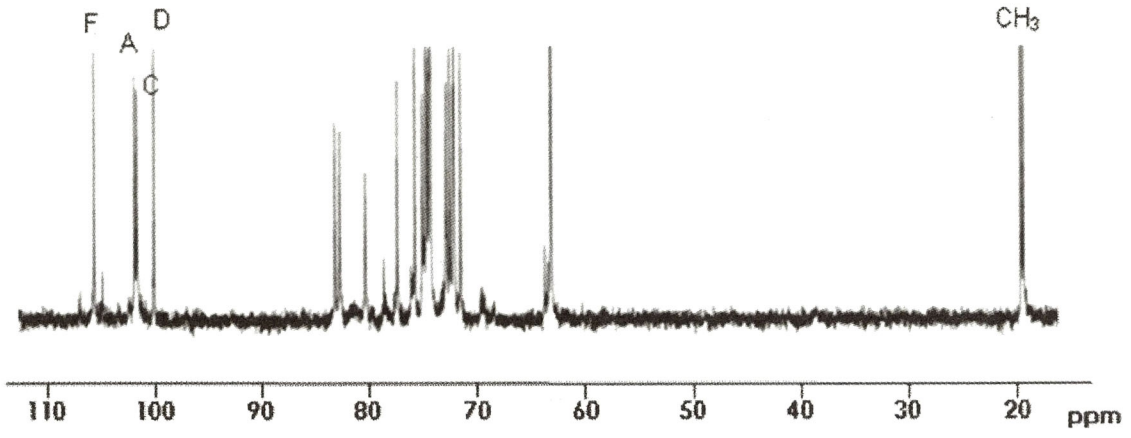


Fig. 3.3 ^{13}C NMR spectrum of Δ*galE* mutant EPS.

Peaks named as in Fig. 3.2.

Table 3.3 ^{13}C chemical shifts of Δ*galE* mutant EPS.

	Chemical shifts(ppm)					
Sugar residue	**C-1**	**C-2**	**C-3**	**C-4**	**C-5**	**C-6**
A → 2)-α-D-Glcp-(1→	101.77	77.50	74.44	74.73	74.80	63.41
C –α-D-Glcp-(1→	101.53	72.33	72.22	75.85	74.80	63.37
D → 2,3)-α-L-Rhap-(1→	100.06	83.27	82.82	73.04	74.38	19.95
F → 3)-α-L-Rhap-(1→	105.60	71.64	80.44	75.15	72.70	19.70

Figure 3.4 shows the inter-residual correlations between the anomeric protons and nearby protons giving information about saccharides sequence in repeating unit. Δ*galE* mutant EPS inter-residue correlations were observed between anomeric proton D and H-2 of A, between anomeric proton C and H-2 of D, and between anomeric proton F and H-3 of D. These correlations, together with methylation analysis results and comparison between wild type CPS and Δ*galE* mutant EPS, allowed us to estimate the monosaccharide sequences of the polysaccharide (the minimum repeating unit), as shown in Fig. 3.5. The repeating unit contained only four monosaccharides instead of six monosaccharides repeating unit of the wild type CPS and Δ*polE* mutant EPS.

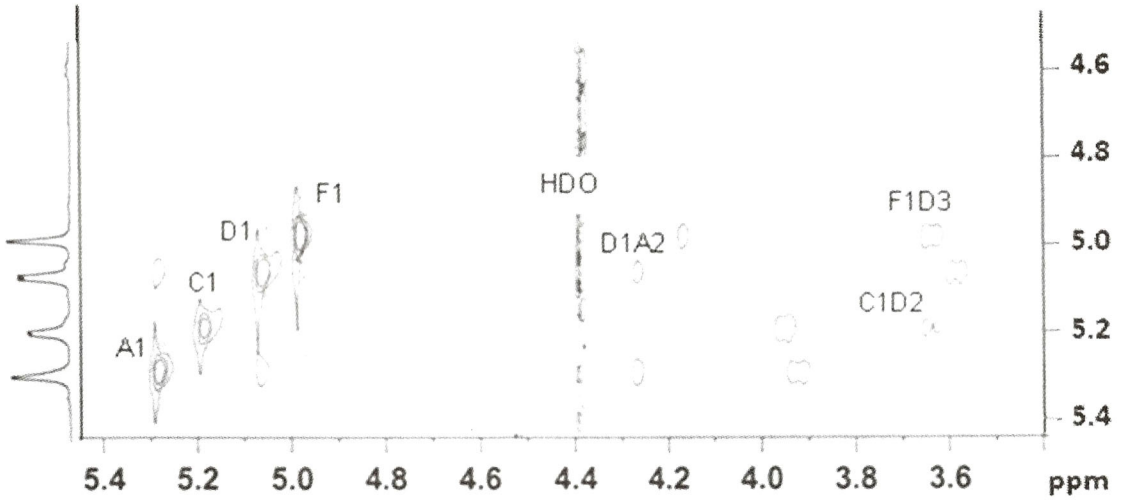


Fig. 3.4 Part of NOESY spectrum of Δ*galE* mutant EPS.

DISCUSSION

The structure of Δ*galE* mutant EPS was identified in this study, and we proposed that the EPS has a branched tetrasaccharide repeating unit composed of 2,3-α-L-rhamnopyranosyl, 3-α-L-rhamnopyranosyl, 2-α-D-glucopyranosyl, and terminal-α-D-glucopyranosyl as a branch (Fig. 3.5). The molecular weight of this EPS was higher (620 KD) than that of the wild type CPS (120 KD) (Moonmangmee et al., 2002b). Monosaccharide compositional analysis revealed that the EPS components were glucose and rhamnose in approximately equimolar ratio using trimethylsilyl (TMS) methylglycoside method, this result is different from that obtained by Deeraksa and co workers (2006), which showed the monosaccharide composition as 1.7:1 glucose:rhamnose. This discrepancy may result from losses in rhamnose during acid hydrolysis and derivatization in alditol acetate method used in that study. Moreover, no acidic residue was detected in this EPS using TMS methylglycoside method which could identify acidic residues.

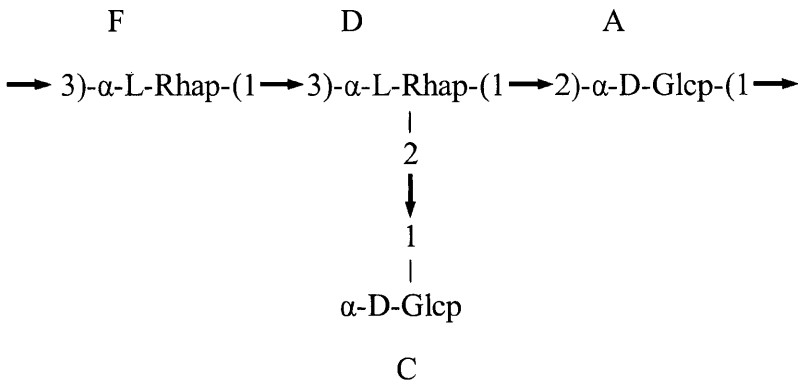


Fig. 3.5 Proposed repeating unit structure of Δ*galE* mutant EPS.

Many minor residues of glucose and rhamnose were detected in methylation linkage analysis, these minor residues were observed in GC-MS spectrum (Fig. 3.1) but not in NMR spectra because of their low abundancy in the EPS. We assume that these minor residues resulted from continuous polymerization in the absence of terminal galactopyranose residue which may terminate the polymerization by forming 2,3,4-rhamnopyranosyl double branch residue in wild type to form CPS (see chapter 2).

Thus, we suggest that this EPS resulted from wild type CPS by missing of terminal galactofuranose and 6-galatopyranose (see chapter 2) due to the *galE* disruption. This could

be supported by the observation that other sugar residues (2,3-rhamnopyranose, 2-glucopyranose, and terminal-glucopyranose) had the same ^{1}H and ^{13}C chemical shifts as the wild type CPS, indicating the same configurations and linkage patterns as in the CPS of the wild type. Additionally the presence of the 3-rhamnopyranosyl residue in Δ*galE* mutant EPS in higher concentration (24.2% mole in Δ*galE* EPS, and 0.34% mole in wild type CPS) could be justified by the loss of terminal galactofuranose branch residue from 2,3-rhamnose branched residue in the repeating unit of the wild type CPS (see chapter 2), which caused decreasing in the branched residue molar ratio from 33.5% in wild type CPS to 23.9% in Δ*galE* EPS (Tables 2.1, 3.1).

Formation of higher molecular weight polysaccharides due to mutations in genes involved in polysaccharide biosynthesis was reported (Xayarath and Yother, 2007). The EPS of Δ*galE* mutant appear to occur by the same mechanism as Δ*polE* mutant EPS which could occur spontaneously due to inactivation of *polE* gene (Deeraksa et al., 2005). As polE may work as β-galactosyltransferase, that could add galactose residues to the repeating unit (see Chapter 2), in Δ*galE* mutant, the inactivation of *galE* results in the absence of the substrate (galactose).

CHAPTER 4

Galactofuranose biosynthesis in *Acetobacter tropicalis* SKU1100

ABSTRACT

Galactofuranose-containing polysaccharides are produced by many bacterial species. These thermodynamically unfavorable furanose sugars are not present in mammalian glycoconjugates, and hence are subject of comprehensive investigations to develop inhibitors of furanose biosynthesis antimicrobial agents. In chapter 2, the galactofuranose residue was shown to be present in *A. tropicalis* SKU1100 wild type CPS. This finding suggests the presence of UDP-galatofuranose mutase (UGM) enzyme encoded by *glf* gene in *A. tropicalis* SKU1100 genome. To test this hypothesis, a *glf* gene was cloned from *A. tropicalis* SKU1100 genome, and the gene disruption was carried out. However, Δ*glf* mutant thus obtained still produced CPS and also formed a pellicle in static culture. Furthermore, ^{13}C NMR of the CPS purified from the mutant confirmed the presence of galactofuranose. The results suggested that another paralogous gene of the tentative *glf* gene may work individually or together with this gene to produce galactofuranose.

INTRODUCTION

Sugars in furanose ring confirmation are found in the surface glycoconjugates of bacteria, fungi and protists. The hexofuranose D-galactofuranose (Galf), the pentofuranose D-arabionofuranose (Araf), and the hexoulose D-fructofuranose (Fruf) are the most occurring furanoses in bacterial polysaccharides (Poulin and Lowary, 2010). Galf is formed from galactopyranose by ring contraction reaction that carried out by UDP-galactopyranose mutase enzyme (UGM). This enzyme is encoded by *glf* gene which has been identified in *Escherichia coli* (Nassau et al., 1996), *Klebsiella pneumoniae* (Köplin et al., 1997), *Mycobacterium tuberculosis* (Weston et al., 1998), and many eukaryotic parasites (Beverley et al., 2005), and also in many bacterial genomes . Moreover, a dual functional UGM, which was renamed as UDP-N-acetylgalactopyranose mutase (UNGM), has been identified in *Campylobacter jejuni* 11168. The enzyme has a relaxed specificity with both UDP-galactose and UDP-N-acetylgalactose as a substrate but higher specificity for UDP-galactose (Poulin et al., 2009). Galf residues are first activated by attaching to UDP group, and then the UDP-Galf is polymerized by galactofuranosyltransferase. In many microorganisms, Galf residues are important for cell viability or play a vital role in cell physiology, depending on whether the Galf-containing surface glycoconjugate is essential for viability or pathogenicity of the microorganism. For example, in *Mycobacterium smegmatis, glf* gene is essential for cell viability because it involves in the biosynthesis of arabinogalacan polysaccharide that connects mycolic acid layer to peptidoglycan layer, and thus could be knocked out only in the presence of two rescue plasmids carrying functional copies of *glf* and a galactofuranosyltransferase gene (Pan et al., 2001). Since mammalians do not have Galf in their glycoconjugates, Galf biosynthesis pathway is an attractive target for research to develop inhibitors of *glf* gene in order to combat diseases such as tuberculosis.

The structural analysis of *A. tropicalis* SKU1100 wild type CPS revealed the presence of Galf residue in the repeating unit (see chapter 2). Hence, since the presence of *glf* gene was expected, cloning and deletion of *glf* was carried out in this chapter.

MATERIALS and METHODS

Bacterial strains and culture conditions

A. tropicalis SKU1100 (R) strain (Deeraksa et al., 2005) was grown at 30°C in potato medium consisting of 5 g of glucose, 10 g of yeast extract, 10 g of polypeptone, 20 g of glycerol and 100 ml of potato extract in 1 L of tap water, or in YPG medium composed of 5 g of yeast extract, 5 g of polypeptone, and 10 g of glycerol in 1 L of tap water. *Escherichia coli* DH5α was grown at 37°C in LB medium consisting of 5 g of yeast extract, 5 g of sodium chloride, and 10 g of polypeptone in 1 L of distilled water, the medium pH was adjusted to 7.5. The transformants of *A. tropiclis* were selected on YPG medium. Antibiotics were added at the concentration of 50 μg/ml as Kanamycin or Ampicillin.

Isolation of chromosomal DNAs

Chromosomal DNA of *A. tropiclis* SKU1100 strains was isolated from cells grown to a late-log phase in 5 ml of potato medium. The cells were harvested by centrifugation at 12,000 rpm for1min at room temperature, and washed first with distilled water, then with sterile 30 mM Tris-HCl buffer (pH 7.9) containing 5 mM EDTA (TE). The obtained pellet was suspended in 400 μl of TE. After 1 mg of lysozyme was added, the cell suspension was incubated at 37 °C for 1 h with gentle shaking. Then, 1/10 volume of 10% SDS was added to the suspension and incubated at 50 °C for 10 min without shaking. Followed by addition of 1/10 volume of 10% cetyltrimethyl ammonium bromide (CTAB) in 0.7 M NaCl and incubation at 65 °C for 10 min without shaking. An equal volume of phenol: chroloform: isoamyl alcohol (25: 24: 1) was then added and vigorously shaken. After shaking, the mixture was centrifuged at 12,000 rpm for 15 min at room temperature. Then the upper layer was collected and 1/10 volume of 3 M sodium acetate was added. As described above, the solution was extracted again with phenol: chroloform: isoamyl alcohol solution. To the upper layer, 2.5-fold volume of cold 100% ethanol was added and DNA was collected by centrifugation at 12,000 rpm for 15 min at 4°C. The DNA was washed with cold 70% ethanol, dried up in vacuum desicator and resuspended in 50 μl of 10 mM Tris-HCl buffer (pH 8.0) containing 1 mM EDTA (pH8.0).

DNA manipulation

Extraction of plasmids from *E. coli* strains were performed with alkali method. PCR was carried out in 25 μl (Go Taq Green Master Mix, Promega Biosciences) or 10 μl (Herculase II, Agilent) of reaction volume using My Cycler PCR machine (BIO-RAD). Agarose gel electrophoresis was performed with 0.8% agarose gel in TBE buffer (89 mM

Tris-HCl (pH 8.0), 89 mM boric acid, 2 mM EDTA), and DNA excised from the gel was purified using a MagExtractor Gel Extraction Kit (TOYOBO). DNA was quantified by Nano Drop ND-1000 spectrophotometer (Applied Biosystems). Primers used in this study are listed in Table 4.1.

Cloning of *glf* gene

To clone *glf* gene primers from the conserved regions of *glf* from *A. pasteurianus* IFO 3283, *Gluconobacter oxydans* 621H, *Agrobacterium radiobacter* K84, *Escherichia coli* SMS-3-5, and *Agrobacterium tumefacieum* str. C58 (Fig. 4.1), galFL and galFR were used (see Table 4.1), PCR was done for 45 cycle of 94°C for 1min, 45°C for 1min, and 72°C for 1min. The 0.9 Kb PCR fragment was ligated to pGEM-T Easy vector (Promega) and subject to sequencing. According to Liu and Whittier (1995) the upstream of the PCR product was cloned using Thermal asymmetric interlaced PCR (TAIL PCR) with 3 specific primers from the known sequence Sp1(-), Sp2(-), and Sp3(-) in addition to arbitrary universal primer AD2 (Table 4.1) using Go Taq polymerase green master mix (Promega). The first TAIL PCR was implemented as one cycle of 92°C for 2min, one cycle of 95°C for 1min, 5 cycle of 94°C for 15sec, 60°C for 1min, and 72 °C for 2 min. Followed by one cycle of 94 °C for 15 sec, 30 °C for 3 min, and 72 °C for 2 min. Then 10 cycles of 94 °C for 5 sec, 44 °C for 1 min, and 72 °C for 2 min, and 12 cycles of 94 °C for 5 sec, 60 °C for 1 min, 72 °C for 2 min, 94 °C for5 sec, 60 °C for 1min, 72 °C for 2 min, 94 °C for 5 sec, 44 °C for 1 min, and 72 °C for 2 min, and finally one cycle of 72 °C for 5 min. The first PCR product was diluted 1000 x times and 1 µl was used as template for the second TAIL PCR which was carried out as 10 cycles of 94 °C for 5 sec, 60 °C for 1 min, 72°C for 2 min, 94 °C for 5 sec, 60 °C for 1min, 72 °C for 2 min, 94 °C for 5 sec, 44 °C for 1min and 72 °C for 2 min, and final cycle of 72 °C for 5 min. The second TAIL PCR was diluted 10 x times and 1µl was used as template for the third TAIL PCR which was done as 20 cycles of 94 °C for 10 sec, 44 °C for 1min, and 72 °C for 2 min, and final cycle of 72 °C for 5 min. The third TAIL PCR product (1 kb) was ligated to pGEM-T Easy vector and sequenced. The downstream of the PCR product was cloned using primers designed from the draft genome sequence *A. tropicalis* glf Ds1(-) and glf Ds2(+). The PCR was done for one cycle of 94 °C for 1 min, and 35 cycles of 94 °C for 30 sec, 45 °C for 30 sec, and 72 °C for 1 min. the PCR product (1 kb) was ligated to pGEM-T Easy vector and sequenced.

Deletion of *glf* gene

In order to delete *glf* gene, a PCR-based method for generating a gene deletion construct was employed (Kuwayama et al., 2002) this method is depending on fusion of upstream and downstream sequences (around 1 kb each) of *glf* gene to kanamycin resistant gene. The upstream and downstream sequences were obtained from the draft genome sequence of *A. tropicalis*. The upstream flanking region was amplified using primers Glf-5(+) and Glf-up-Km(-) which was designed to have an overlapping sequence with the kanamycin resistant gene. The PCR amplification was done with Herculase II DNA polymerase, the reaction mixture (10 µl) contained 2 µl 5x Herculase II buffer, 1 µl Glf-5(+) primer 1 µl, Glf-up-km(-) primer 1 µl, dNTPs (2 mM each) 1.25 µl, genomic DNA 1 µl (about 10 ng), 3.65 µl of sterile distilled water, and 0.1 µl Herculase II. PCR was implemented as one cycle of 94°C for 1 min, and 25 cycles of 94 °C for 30 sec, 55 °C for 30 sec, 68 °C for 1 min, plus one cycle of 68 °C for 5min. For downstream amplification, primers Glf-3-2(-) and Km-glf-down (+) which was also designed to have an overlapping sequence with the kanamycin resistant gene were used. The PCR was done as mentioned above. While kanamycin resistant gene was amplified from the plasmid pTkm (Yoshida et al., 2003) using primers Glf-up-km(+) and Km-glf-down(-), both of which have overlapping ends with upstream and downstream sequences of *glf* gene, respectively. PCR amplification was done as mentioned above. The 3 PCR products were purified by MagExtractor gel extraction kit, and fusion PCR was carried out in 20 µl reactions containing 4 µl 5x Herculase II buffer, 2 µl of Glf-5(+) primer, 2 µl of Glf-3-2(-) primer, 9.3 µl of DNA fragments (1 µl of kanamycin resistant gene, 3 µl of 5' region of *gfl*, and 5.3 µl of 3' region of *glf*). The fusion PCR cycling was: one cycle of 94 °C for 1 min; 25 cycles of 94 °C for 30 sec, 55 °C for 30 sec, and 68 °C for 3 min; and one cycle of 68 °C for 5 min. The resulting 2.9 kb fusion PCR product was purified from agarose gel by MagExtractor gel extraction kit and 200 ng were electroporated into *A. torpicalis* SKU1100 (R) strain. The transformant strains were selected on YPG medium containing kanamycin. The deletion of the target gene was confirmed by DNA sequencing.

Table 4.1 The sequence of the primers used in this study.

Primer	Sequence (5' – 3')
galF L	CATAAGTACGGCCCGCATATTTTC
galF R	CATATCCAGCTAACGGTACGTGCC
Sp1(-)	GGAAGCGTATAGACACGCCC
Sp2(-)	CGCCAAAACGCTGCACGTAG
Sp3(-)	ATCCCGTTCGGTATGGCAGG
AD2	NGTCGASWGANAWGAA[a]
glf DS1(-)	TGGCAAAATGGTAGGCAAGC
glf DS2(+)	CTGCCATACCGAACGGGATG
Glf-5(+)	TCTGAACACGGATGCGGAAG
Glf-3-2(-)	AACCGCTGCTCATGAATTCC
Glf-up-km(+)	CCACATTTTCCATACTGCAGCGAACCGGAATTGCCAGCTG
Glf-up-km(-)	CAGCTGGCAATTCCGGTTCGCTGCAGTATGGAAAATGTGG
Km-glf-down(+)	CTTGACGAGTTCTTCTGAGACTGCAGCGTTACAGGCGAAAG
Km-glf-down(-)	CTTTCGCCTGTAACGCTGCAGTCTCAGAAGAACTCGTCAAG

[a] N: any nucleotide, S: G or C, W: A or T.

```
A.pasteurianus  CGAAACCGGCGTTATGGTGCATAAATATGGTCCCCATATTTTCCACACTG 184
G.oxydans       CAAGACCGGTGTGATGGTGCATCGCTATGGGCCTCACATTTTCCATACTG 184
Ag.radiobacter  CGA--CCGGTGTCATGGTGCATGTCTACGGCCCGCATATTTTCCACACGG 196
Ag.tumefacieum  CGAAACCGGCGTCATGGTGCATGTTTACGGTCCGCATATTTTCCATACCG 196
E.coli          AGA--CTGATGTCATGGTTCATGTCTACGGTCCGCATATTTTCCATACGG 193
                  *  * *  ** ***** ***   ** ** ** ** ******** ** *

A.pasteurianus  CAGATGAGCGGACATGGGCCTATGTGCAAAAATTTG-GCACGTTCAAGCC 233
G.oxydans       CCGACGAACGCGCTTGGAGCTACATTCAGCGTTTTG-GAAAATTCCGGCC 233
Ag.radiobacter  ACGACAGGGAAGTCTGGGACTATGTGAA-CCTCTTTCAGACGTTTATGCC 245
Ag.tumefacieum  ACGACCGGGAAGTCTGGGACTACGTGAA-CAGCTTCCAGACCTTCATGCC 245
E.coli          ATAATGAAACCGTCTGGAACTATGTCAATCAGTATGCGGAAATG-GTGCC 242
                   *          ***  ***  *  *      *    *  *    ***

A.pasteurianus  TTATGTAAACCGCGTGAAGGCCATTTCTAACGAGCACGTTTACACGCTGC 283
G.oxydans       TTATATCAACCGTGTGAAGGCAATTTCCCAAGGGCGCATTTACACGTTGC 283
Ag.radiobacter  CTACAAGAACCGCGTCAAGACGACGAGTTCGGGCCAAGTCTATTCCTTGC 295
Ag.tumefacieum  CTACAAGAACCGCGTCAAAACGACGAGCGGTGAGCAGGTCTACTCCCTTC 295
E.coli          CTATGTGAATAGAGTAAAAGCTACTGTTAATGGTCAGGTTTTTTCACTGC 292
                 **    **  * ** **  * *        *  *   * *   *  * *

A.pasteurianus  CCGTTAACCTGCTGACCATTAACCAGTTTTTCGGCACAACGATGGGGCCG 333
G.oxydans       CTGTCAACCTTCTGACGATTAACCAGTTCTTCGGTACGGTTATGTCACCT 333
Ag.radiobacter  CCATCAATCTTCACACGATCAATCAGTTTTTCGGCAAGACGCTCAGACCG
Ag.tumefacieum  CCGTCAATCTGCACACGATCAACCAGTTCTTCGGCAAGAATTTCCGGCCC 345
E.coli          CGATTAATCTGCATACTATTAATCAGTTCTTTGCTAAAACCTGTTCTCCT 342
                *  * ** ** *  ** ** ** ***** ** *  *          **

A.pasteurianus  AAGGAAGCACGTGCTTTTATTGAAAGTAAGGCAGATAAAAGCATTACCAA 383
G.oxydans       GTAGAGGCGCGGGCGTTTATTGAAAAGAAGGCCGATAAGTCGATCAAGGA 383
Ag.radiobacter  GACGAAGCGCGTGCTTTCATCGAGGAGCGTGCCGACAAAAGCATAACGGA 395
Ag.tumefacieum  GATGAAGCGCGAACCTTCATCGAGGACAAGGCCGATAAATCGATCACCGA 395
E.coli          GATGAAGCTCGCGCGCTCATTAGCGAAAAAGGTGATAGTTCCATTCTGGA 392
                   ** ** **  *  * **          *  ** *     **     *

A.pasteurianus  CCCTCAGTCCTTTGAAGAACAGGCACTGAGCATGATTGGCCCGGAACTGT 433
G.oxydans       TCCAAAAAATTTCGAAGAGCAGGCGCTCTCGATGATTGGTCCGGAGCTAT 433
Ag.radiobacter  TCCGAAGACCTTCGAGGAGCAGGCGCTGCGCTTCGTCGGCAATGACCTGT 445
Ag.tumefacieum  TCCACAGACCTTCGAGGAGCAGGCGCTGCGCTTCGTCGGCCGCGATCTCT 445
E.coli          ACCGAAGACTTTTGAAGAGCAGGCCTTGCGCTTCATTGGTAAAGAATTAT 442
                 **  *    ** ** ** *****  *     *  * **    **  * *

A.pasteurianus   ATCGGGCGTTCTTTTATGGATATACGCGCAAGCAGTGGGGGCTGGA-GCC 482
G.oxydans        ACAAGGCGTTTTTCTACGGATATACGCGAAAGCAGTGGGGGCTGGA-GCC 482
Ag.radiobacter   ATGAAGCGTTCCTCAAGGGATATACCCAGAAGCAATGGGGT-TGCGCGCC 494
Ag.tumefacieum   ATGAGGCCTTCTTCAAGGGATATACCGAAAAGCAATGGGGC-TGCTCGCC 494
E.coli           ACGAAGCGTTCTTCAAAGGCTACACCATAAAACAGTGGGGGATGCA-GCC 491
                 *    ** **  *  * ** ** **    ** ** *****  **   ***

A.pasteurianus   TACGCAGCTTCCCGCATCTATTCTCAAGCGCCTGCCTTTGCGCTTCAATT 532
G.oxydans        GACAGAACTGCCTGCGTCCATTCTCAAAAGATTGCCGCTGCGCTTTAATT 532
Ag.radiobacter   GACCGAGCTTCCAGCCTCTATCCTGAAACGGCTTCCGGTCCGCTTCAACT 544
Ag.tumefacieum   GACGGAACTGCCGGCCTCTATCCTGAAACGCCTGCCCGTGCGCTTCAATT 544
E.coli           TTCTGAGCTTCCGGCATCTATACTCAAACGTTTGCCTGTGCGTTTCAATT 541
                   *  * ** ** ** ** ** ** **  *  * **  * ** ** ** *

A.pasteurianus   ATGATGATAACTACTTCAACCATCCGTATCAGGGGATGCCTGAAGAAGGT 582
G.oxydans        ATGATGACAATTACTTTAATCACCCGTATCAGGGGATGCCGGAAGACGGG 582
Ag.radiobacter   ACGACGACAACTATTTCTTCCATAAATACCAGGGCATGCCGGAAAACGGC 594
Ag.tumefacieum   ACGACGACAATTACTTCTTCCACAAATATCAGGGCATGCCGGAAAGCGGT 594
E.coli           ATGACGATAATTATTTCAATCATAAATTCCAGGGCATGCCGAAGTTAGGC 591
                 * ** ** ** ** **    **    *  ***** *****  *    **

A.pasteurianus   TATACAGCCATTGTACA-GAATATCCTGAATGTTGAGGGTATTGATGTGC 631
G.oxydans        TATTCAGCGATTGTAGA-AAATATTCTGAAGTCTGAAAATATTGAGATTC 631
Ag.radiobacter   TATACCGACATGGTCGG-GCGAATTCTCGATCACACGAATATAGAAGTCC 643
Ag.tumefacieum   TATACTGACATGATCGA-GCGCATCCTCGATCACCCCAATATCACGGTAA 643
E.coli           TATACCCATATGATTGAAGCG-ATCGCCGATCATGAAAATATCACTCTGC 640
                 *** *    **  *        **     *          ***     *
A.pasteurianus   GGTTGGGCTGTTCTTTTGAAAGCCTGGAA----GAAGATTTTCAG-CACG 676
G.oxydans        GCCTCAATACAAAATTCGAGGATGTTTCG----GAAGAATTTCAG-CATA 676
Ag.radiobacter   ACTTGGAAACCAGGTTTCACAGCGGTCTA--GCCAAGGACTTCGATCACA 691
```

```
Ag.tumefacieum   AACTCGGCACGCGCTTTGACCGCGCCGAA--GCGCCGGCCACCGGACACG 691
E.coli           AATT-ACAGCGTGAGTTTGCGGCTGAAAATCGCGAAAGTTAT-GATCATG 688
                    *           *                              **

A.pasteurianus   TATTCTATACCGGCCCGATTGATCGGTATTTTGGTTTCCGTCTGGGCCGC 726
G.oxydans        TCTTCTATACTGGGCCGATAGATCGTTATTTCTCTTTTGATCTCGGGCGT 726
Ag.radiobacter   TATTCTATTCCGGTCCCCTCGACGGCTATTTCGATTATGAATATGGCCGT 741
Ag.tumefacieum   TCTTTTACTCCGGTCCGCTCGATGGCTATTTCGATTTCGAATTCGGCCGC 741
E.coli           TATTTTATAGCGGGCCGTTAGACGCCTTCTATTCATACCAGCATGGGCGT 738
                 * ** **    ** **  * **    *  *     *        ** **

A.pasteurianus   TTGGGTTACAGGACGTTGGATTTTGAACGCTTCGTAGATGATGGTGATCA 776
G.oxydans        CTGGGATATCGTACGCTCGATTTTGAGCGTATCGATGGAGATGGTGATTA 776
Ag.radiobacter   TTGGGATATCGCACGCTGGATTTCGAGCGTTTTACCCATGATGGTGACTA 791
Ag.tumefacieum   CTGGCGTACCGGACGCTCGACTTCGAACGGTTCACCTATGACGGCGACTA 791
E.coli           CTTGGCTACCGTACTCTGGACTTTGAGCGATTTACCTGGCAAGGTGATTA 788
                  * *  **  * **  * ** ** ** **  *        * ** **  *

A.pasteurianus   TCAGGGCACAGCCGTTATCAACTATTGCAATGAAGAAGTGCCTTACACGC 826
G.oxydans        TCAGGGCGCAGCAGTAATTAATTACTGCGATCAGGAGGTGCCGTTCACGC 826
Ag.radiobacter   TCAAGGGTGCGCAGTCATGAACTATGGCGATGTCTCCGTGCCCTATACGC 841
Ag.tumefacieum   CCAGGGCTGCGCGGTGATGAATTACGGCGACGTGTCGGTACCCTTTACCC 841
E.coli           CCAGGGCTGCGCGGTCATGAACTATTGCTCAGTTGATGTCCCGTATACGC 838
                  ** **    ** ** ** ** **  **         ** ** *  ** *

A.pasteurianus   GTATTTCTGAACACAAGCACTTTGCCCCATGGGAGCAGGACAAATTCTCC 876
G.oxydans        GTATCTCAGAGCATAAACATTTTGCTCCATGGGA----AGCAGATTCCCT 872
Ag.radiobacter   GCGTGACGGAGCATAAGCATTTCTCCCCGTGGGA--AGACCATGCCGGCT 889
Ag.tumefacieum   GCATCACCGAACACAAGCATTTTTCGCCGTGGGA--AGACCATGCCGGCT 889
E.coli           GCATTACCGAACATAAATATTTTTCCCCGTGGGA--AAACCACGAAGGCT 886
                 *  *  * ** ** **  * **  * ** *****      **      *

A.pasteurianus   ----AAAACTGTTTGTTTCCGTGAATACAGCCGC-CTGGCAGGTGAAAAC 921
G.oxydans        GCAAAAGACAGTTGCTTTCAGAGAATACAGCCGT-CTGGCGGAACCTAAA 921
Ag.radiobacter   --------CAGTCTGCTATCGCGAGTTCTCCAGGGCCTGC-GAACCCGGC 930
Ag.tumefacieum   --------CCGTCTGCTACCGGGAGTTTTCCCGCGAATGC-GGCCCACAG 930
E.coli           --------CCGTATGTTACAAAGAATATAGTCGCGCCTGC-GGTGAAAAT 927
                        * **    *     ** *      *     ** *

A.pasteurianus   GATATTCCCTACTACCCCATCCGCCTGGTGAACGAAAAGAAGATGCTCGA 971
G.oxydans        GACGTTCCGTATTATCCAATTCGCCTCGCTAATGAAAAAACCATGCTCGA 971
Ag.radiobacter   GATGTTCCCTACTATCCGATCCGCCTGATGGAGGAAAAGGAACAGCTCGC 980
Ag.tumefacieum   GACATTCCCTATTATCCGATCCGGCTGGTCGAGGATAAAGAACAGCTCGC 980
E.coli           GATATTCCTTACTATCCAATCCGCCAGATGGGCGAGATGGCTCTGCTGGA 977
                 **  **** ** ** ** ** ** *        ** *       *** *

A.pasteurianus  AAGCTATATCGAGCTGGCACGTGCAGAAAAAGGCGTGTCCTTTATGGGGC 1021
G.oxydans       GAGCTACATTACTCGGGCCCGAAATACGGCAGGCGTATCTTTTCTCGGCC 1021
Ag.radiobacter  GCACTATGTCACGCGAGCCCTCCAGGAGCCCTCCGTCACCTTTGTCGGCC 1030
Ag.tumefacieum  CGACTATGTGGCGCGGGCCGAACGGGAAGCGTCGGTCACCTTTGTCGGAC 1030
E.coli          TAAATATCTATCACTGGCGGAAGGTGAGAAAAATATTACTTTCGTCGGAC 1027
                    **  *    *  **                 *  * **  * ** *

A.pasteurianus  GTTTGGGCACGTATCGTTATCTGGATATGGATGTGACCATCAAGGAAGCT 1071
G.oxydans       GGCTTGGGACATACCGGTACTTGGATATGGACGTGACGATTTCTGAAGCT 1071
Ag.radiobacter  GGCTCGGGACATACCGCTATCTCGATATGGACGTCACGATCCGCGAGGCG 1080
Ag.tumefacieum  GGCTCGGGACCTACCGTTATCTCGACATGGATGCGACCATCCGCGAAGCG 1080
E.coli          GCTTAGGGACTTATCGCTATCTTGATATGGATGTAACGATTGCAGAAGCT 1077
                *  *  ** ** ** ** **  * ** *** ** *  ** **    ** **
```

Fig. 4.1 Nucleotide sequence alignments of *glf* from *A. pasteurianus* IFO 3283 (A.pasteurianus), *Gluconobacter oxydans* 621H (G.oxydans), *Agrobacterium radiobacter* K84 (Ag.radiobacter), *Agrobacterium tumefacieum* (Ag. tumefacieum), and *Escherichia coli* SMS-3-5 (E.coli).

Asterisks indicates identical nucleotides, arrows indicate primers locations. In primers design, when nucleotides are not identical *A. pasteurianus* IFO 3283 nucleotides and amino acid codons were used.

Nucleotide sequence analysis

The DNA sequencing was carried out after ligating the PCR products to pGEM-T Easy vector as described above, using an ABI PRISM310 (PE Biosystems), and the data were analyzed using the GENETYX-MAC program (Software development Co., Ltd). Similarity search analysis and alignment were performed using BLAST (Atschul *et al*, 1990) and CLUSTAL W (www.ebi.ac.uk/clustalw), respectively.

Purification of Δ*glf* mutant CPS

The CPS of Δ*glf* mutant was purified from cells as mentioned in Chapter 2. In brief, 10% of the seed culture was inoculated into 1 L of potato medium and incubated at 30°C with shaking for 30 h. The cells were then collected, washed 2 times with 50 mM phosphate buffer (pH 6.5), and resuspended in the same buffer (1 g cells/10 ml buffer). The suspension was ultra-sonicated using sonicating microprobe (TP-040, 3mm diameter, TOMY TECH, INC) for 20 min, centrifuged, and ultra-centrifuged to remove cell debris. DNase (50μg/ml) was added to the supernatant and it was incubated at 37°C overnight, followed by an additional overnight incubation with 100 μg/ml proteinase K at 37°C. The suspension was then subjected to dialysis against 25 mM Tris-HCl buffer (pH 8.5) overnight. After centrifugation to remove precipitate, the supernatant was applied to a DEAE-cellulose column and polysaccharides were eluted with 25 mM Tris-HCl (pH 8.5). Polysaccharide fractions were determined by phenol-sulfuric acid assay (Dubois et al., 1956), pooled, ultra-centrifuged, and precipitated with 2 volumes of cold ethanol. The precipitated polysaccharide was then dissolved in 0.1M NaCl and applied to a Superdex S-200 column. The polysaccharide fractions were pooled and precipitated with 2 volumes of cold ethanol.

NMR analysis

The purified CPS was dissolved in D_2O and lyophilized 3 times to reduce the residual water. ^{13}C NMR analysis was done in a Bruker 400 MHz spectrometer with 5 mm sample probe, which was kept at 60°C. Standard pulse sequences from Bruker were used to obtain $1D^{13}C$ NMR spectra. Chemical shifts were reported in ppm down field from TSP internal standard.

```
Glf-A.t         MS---FCVVGAGFSGAIIARHLAERGHKVTVVDERPHIGGTCHTERDAETGVMVHKYGPH 57
Glf-A.p         MT---FCVVGAGFSGAVIARHLAERGHKVVVVDERAHIGGNCYTERDTETGVMVHKYGPH 57
Glf-G.o         MR---FCIVGAGFSGAIIARHLSEKGHKVLLLDERSHIAGNCHTERDDKTGVMVHRYGPH 57
Glf-A.r         MR---FCIVGAGFSGAIIARHLSEKGHKVLLLDERSHIAGNCHTERDDKTGVMVHRYGPH 57
Glf-E.c         MKRNNILIVGAGFSGVVIARQLAEQGHKVKIIDQRDHIGGNSYDTRDPQTDVMVHVYGPH 60
                *    : :*******.:***:*:*:**** ::*:* **.*..:  ** :*.**** ****

Glf-A.t         IFHTADQRAWDYVQRFGAFKPYVNRVKAISQGRVYTLPVNLLTINQFFGTTMGPKEARAF 117
Glf-A.p         IFHTADERTWAYVQKFGTFKPYVNRVKAISNEHVYTLPVNLLTINQFFGTTMGPKEARAF 117
Glf-G.o         IFHTADERAWSYIQRFGKFRPYINRVKAISQGRIYTLPVNLLTINQFFGTVMSPVEARAF 117
Glf-A.r         IFHTADERAWSYIQRFGKFRPYINRVKAISQGRIYTLPVNLLTINQFFGTVMSPVEARAF 117
Glf-E.c         IFHTDNETVWNYVNQYAEMVPYVNRVKATVNGQVFSLPINLHTINQFFAKTCSPDEARAL 120
                **** :: .* *:::::. : **:*****  : ::::**:** ******... .* ****:

Glf-A.t         IEGKADKTITNPRSFEEQALSMIGPELYRAFFYGYTRKQWGLEPTELPASILKRLPLRFN 177
Glf-A.p         IESKADKSITNPQSFEEQALSMIGPELYRAFFYGYTRKQWGLEPTQLPASILKRLPLRFN 177
Glf-G.o         IEKKADKSIKDPKNFEEQALSMIGPELYKAFFYGYTRKQWGLEPTELPASILKRLPLRFN 177
Glf-A.r         IEKKADKSIKDPKNFEEQALSMIGPELYKAFFYGYTRKQWGLEPTELPASILKRLPLRFN 177
Glf-E.c         ISEKGDSSILEPKTFEEQALRFIGKELYEAFFKGYTIKQWGMQPSELPASILKRLPVRFN 180
                *. *.*.:* :*:.****** :** ***.*** *** ****::*:**********:***

Glf-A.t         YDDNYFNHPYHGMPEEGYTAIVQNILNIEGIDLRLGCSFESLE-EKFQHVFYTGPIDRFF 236
Glf-A.p         YDDNYFNHPYQGMPEEGYTAIVQNILNVEGIDVRLGCSFESLE-EDFQHVFYTGPIDRYF 236
Glf-G.o         YDDNYFNHPYQGMPEDGYSAIVENILKSENIEIRLNTKFEDVS-EEFQHIFYTGPIDRYF 236
Glf-A.r         YDDNYFNHPYQGMPEDGYSAIVENILKSENIEIRLNTKFEDVS-EEFQHIFYTGPIDRYF 236
Glf-E.c         YDDNYFNHKFQGMPKLGYTHMIEAIADHENITLQLQREFAAENRESYDHVFYSGPLDAFY 240
                ******** ::***: **: ::: * . *.* ::*  .*   . *.::*:**:**:* ::

Glf-A.t         QFKLGRLGYRTLDFERFVDEGDHQGTAVINYCDENVPFTRISEHKHFAPWEQDKFSKTAC 296
Glf-A.p         GFRLGRLGYRTLDFERFVDDGDHQGTAVINYCNEEVPYTRISEHKHFAPWEQDKFSKTVC 296
Glf-G.o         SFDLGRLGYRTLDFERIDGDGDYQGAAVINYCDQEVPFTRISEHKHFAPWEADSLQKTVA 296
Glf-A.r         SFDLGRLGYRTLDFERIDGDGDYQGAAVINYCDQEVPFTRISEHKHFAPWEADSLQKTVA 296
Glf-E.c         SYQHGRLGYRTLDFERFTWQGDYQGCAVMNYCSVDVPYTRITEHKYFSPWEN--HEGSVC 298
                 :  ************:  :**:** **:***. :**:***:***:*:***    . :..

Glf-A.t         FREYSRLAGEKDVPYYPIRLVNEKQMLESYIGLARAEKGVSFLGRLGTYRYLDMDVTITE 356
Glf-A.p         FREYSRLAGENDIPYYPIRLVNEKKMLESYIELARAEKGVSFMGRLGTYRYLDMDVTIKE 356
Glf-G.o         FREYSRLAEPKDVPYYPIRLANEKTMLESYITRARNTAGVSFLGRLGTYRYLDMDVTISE 356
Glf-A.r         FREYSRLAEPKDVPYYPIRLANEKTMLESYITRARNTAGVSFLGRLGTYRYLDMDVTISE 356
Glf-E.c         YKEYSRACGENDIPYYPIRQMGEMALLDKYLSLAEGEKNITFVGRLGTYRYLDMDVTIAE 358
                ::**** .  :*:****** .*  :*:.*:  *.   .::*:*************** *

Glf-A.t         ALAACDRIDAALQAKELDLPAFFIDPS 383
Glf-A.p         ALAACDVIDTAMQASELSLPAFFVDPA 383
Glf-G.o         ALKACDQIDQLLAANAP-LPAFFVDPS 382
Glf-A.r         ALKACDQIDQLLAANAP-LPAFFVDPS 382
Glf-E.c         ALKTADKYLSSLSNDET-MPVFVADVR 384
                ** :.*     :  .   :*.*. *
```

Fig. 4.2 Alignment of the deduced amino acid sequences of *glf* from *A. tropicalis* (A.t), *A. pasteurianus* IFO 3283 (A.p), *Gluconobacter oxydans* 621H (G.o), *Agrobacterium radiobacter* K84 (A.r), and *Escherichia coli* SMS-3-5 (E.c), done with ClustalW (http://www.ebi.uk/clustalW).

Asterisks indicate amino acids identical in all sequences, while dots represent similar amino acids. Numbers showed amino acid sequence position. – indicates space introduced by ClustalW alignment.

RESULTS

Cloning of a *glf* gene from *A. tropicalis*

The repeating unit structure of *A. tropicalis* SKU1100 CPS was found to contain Galf residue (see Chapter 2). To identify a *glf* gene, 0.9 kb internal PCR fragment of the gene was obtained by PCR using primers designed based on the conserved regions of *glf* from *A. pasteurianus* IFO 3283, *Gluconobacter oxydans* 621H, *Agrobacterium radiobacter* K84, *Escherichia coli* SMS-3-5, and *Agrobacterium tumefacieum* str. C58 (Fig. 4.1). The PCR product yielded 85% identity to *glf* from *A. pasteurianus*. Thus, TAIL PCR and PCR with primers designed from the draft genome sequence of *A. tropicalis* were used to obtain the gene upstream and downstream regions, respectively. Sequence analysis of the DNA fragments revealed an open reading frame (ORF) of 1156 bp. A promoter-like sequences -35 (5'-TTAAAT-3') and -10 (5'-GAATAAGCT-3') were found, while no terminator was found downstream, indicating that the ORF may form an operon with other genes downstream. Upstream of this gene, a small ORF with no similarity to known sequences and ORF having similarity to leucyl/cytosol amino peptidase were found. While, a hypothetical protein was found downstream (Fig. 4.3).

A

B

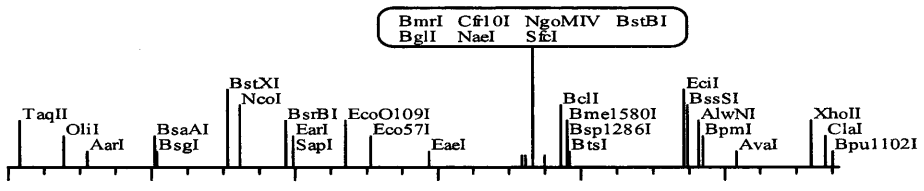


Fig 4.3 Gene organization of *glf* (A), and restriction map indicating enzymes that cut one time (B).

Filled triangle indicates the promoter like sequence.

Effect of *glf* deletion on pellicle formation

Glf gene is involved in UDP-galactofuranose biosynthesis from UDP-galactopyranose, and thus the deletion of *glf* may affect CPS biosynthesis and pellicle formation. To identify the function of the cloned *glf* gene, PCR-based gene deletion method was employed (Fig. 4.3), and the *glf* deletion was confirmed by sequencing. The resultant Δ*glf* mutant still produced pellicle in static culture. To further check the effect of *glf* deletion, CPS of the mutant was purified and subjected to monosaccharide composition and NMR analysis. The CPS showed the same monosaccharide composition, composed of glucose, galactose, and rhamnose, as the wild type CPS. Moreover, ^{13}C NMR analysis showed six anomeric carbon signals: A, B, C, D, E, and G (Fig. 4.4) which had the same chemical shifts of 102.35, 104.31, 101.07, 100.76, 110.41, and 107.4 as wild type CPS. The presence of Galf signal (E at 110.41 ppm) indicated that the Δ*glf* mutant still produces Galf in its CPS. Since deletion of *glf* gene abolished CPS production in *Cambylobacter jejuni* (Poulin et al., 2009), the same effect was expected to occur in *A. tropiclis* SKU1100 or at least Galf residue to be missed from the CPS structure. However, since Δ*glf* mutant did not lose the CPS or Galf residue, the presence of another copy of *glf* gene was expected, and 2 another similar sequences were found with the BLAST search in the draft genome of *A. tropicalis*.

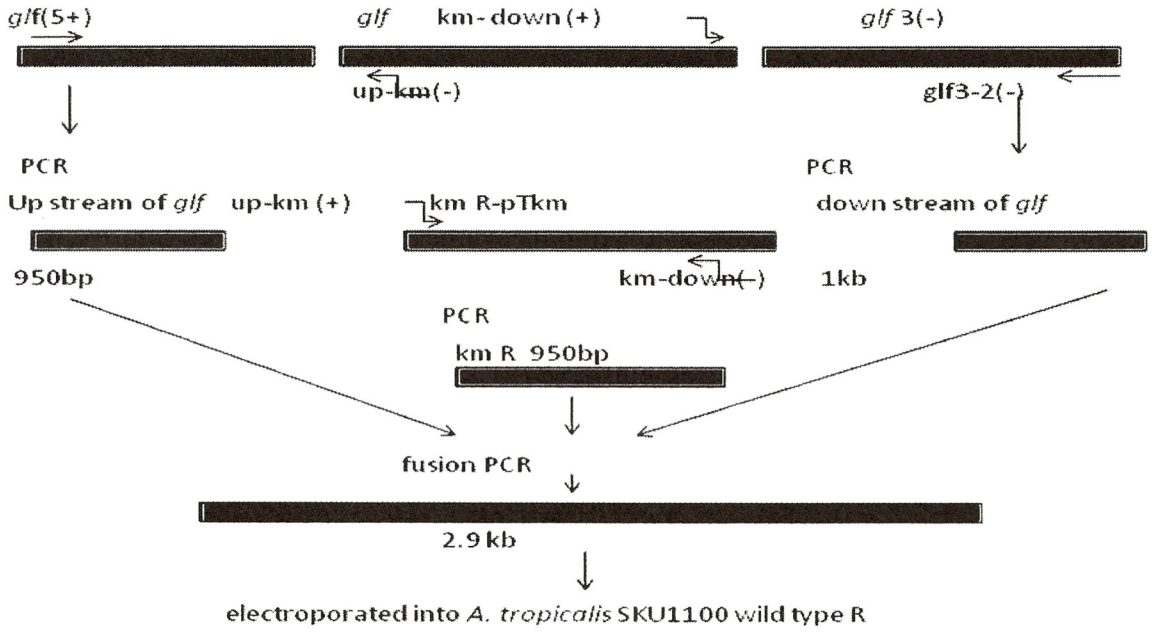


Fig. 4.4 Construction of *glf* gene deletion fusion PCR fragment. Arrows indicate primes locations.

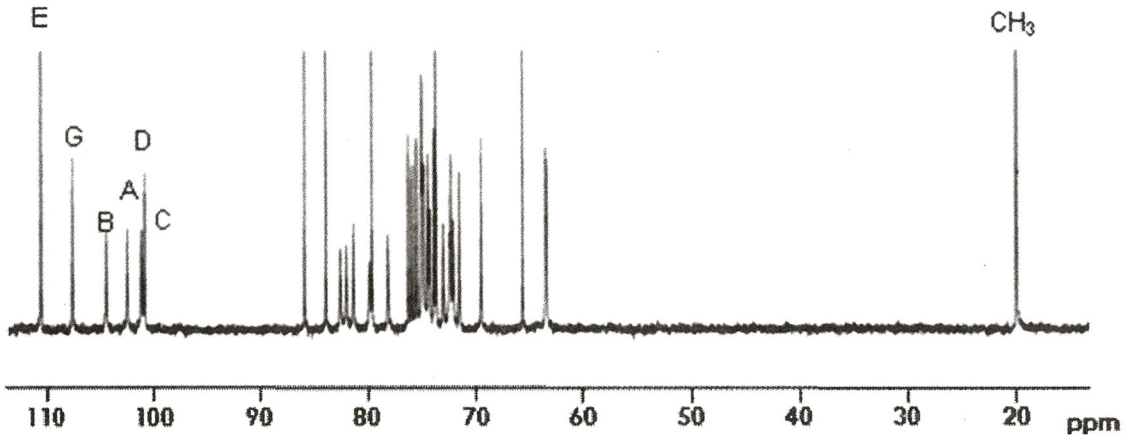


Fig. 4.5 ^{13}C NMR spectrum of Δ*glf* mutant CPS. The spectrum was obtained in D_2O at 60°C with TSP as internal standard.

The signal label is the same as in the wild type CPS, signal E (110.41 ppm) was identified as Galf residue (see Chapter 2).

DISCUSSION

Acetobacter tropicalis SKU1100 produces CPS which is important for pellicle formation and growth in static culture. The pellicle CPS was composed of glucose, galactose, and rhamnose, and the structural characterization of the CPS revealed the presence of Galf sugar residue as well as galactopyranose in the repeating unit. Deeraksa et al., (2006) investigated the biosynthesis of galactose in *Acetobacter tropicalis* SKU1100, and found that *galE* gene (UDP-galactose 4-epimerase) is responsible for galactose biosynthesis in this strain. However, *galE* did not form an operon with other polysaccharide biosynthesis genes such as *glf* gene, which is presumably involved in biosynthesis of Galf from galactopyranose.

In this study, a gene encoding UDP-galactopyranose mutase, *glf,* was identified by PCR using degenerate primers based on conserved regions of *glf* from many bacteria. The deduced amino acids sequence showed a high homology to *glf* from *A. pasteurianus* IFO 3283 and other related bacteria. A promoter like sequence was found in the upstream area, while no transcriptional terminator was detected downstream of the gene, implying that the gene may be the first one in an operon. The *glf* gene was flanking a small ORF with no homology to known genes in BLAST search, and upstream of hypothetical protein that may form an operon with *glf.* The *glf* gene was found to be located in *rfb* region in *E. coli* K12

(Nassau et al., 1996) and in the O-polysaccharide gene cluster in *Klebsiella pneumoniae* (Guan et al., 2001). In contrast, *Mycobacterium semgmatis glf* gene was found to be the first in an operon containing galactofuranosyltransferase and other 3 ORFs with unknown function (Pan et al., 2001).

Although the gene was deleted, the Δ*glf* mutant strain still produced pellicle and CPS as wild type strain. Moreover, ^{13}C NMR spectrum of the polysaccharide purified from the mutant strain confirmed the presence of Galf residue in the CPS. It is noteworthy that Galf production is not essential for cell viability in *A. tropicalis* because the Δ*galE* mutant did not produce both Galf and galactopyranose (see Chapter 3). Thus, the gene identified in this study may have another paralogous gene that works individually to produce Galf. This assumption could be supported by the observation that there are some homologous to the cloned *glf* in more than one contigs of the draft genome sequence of *A. tropicalis*. Gene duplication is common among many genomes, and this could be happen via the processing of unequal crossing-over, unequal sister chromatid exchange, and DNA amplifications or replication slippage (Brown, 1999). In conclusion, since the deletion of *glf* did not abolish Galf production, another copy of the gene is expected to exist. Hence, further investigation is needed to identify this paralog and elucidate its role in Galf biosynthesis.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Boil. 215*: 403-410.

Altschul, S. F., Madden, T.L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25*: 3389-3402.

Azuma, Y., Hosoyama, A., Matsutani, M., Furuya, N., Horikawa, H., Harada, T., Hirakawa, H., Kuhara, S., Matsushita, K., Fujita, N., & Shirai, M. (2009). Whole-genome analyses reveal genetic instability of *Acetobacter pasteurianus*. *Nucleic Acids Res. 37*: 5768-5783.

Bae, E., & Phillips Jr., G.N. (2004). Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. *J. Biol. Chem. 279*: 28202-28208.

Bae, E., & Phillips Jr., G.N. (2005). Identifying and engineering ion pairs in adenylate kinases. Insights from molecular dynamics simulations of thermophilic and mesophilic homologues. *J. Biol. Chem. 280*: 30943-30948.

Beverley, S. M., Owens, K. L., Showalter, M., Griffith, C. L., Doering, T. L., Jones,V. C., & McNeil, M. R. (2005). Eukaryotic UDP-galatopyranose mutase (GLF Gene) in microbial and metazol pathogens. *Eukaryotic Cell, 4*: 1147-1154.

Brown, T. A. (1999). Genomes. John Willey and sons, New York.

Ciucanu, I., & Kerek, F. (1984). A simple and rapid method for the permethylation of carbohydrates. *Carbohydr. Res., 131*: 209-217.

Cleenwerck, I., Vandemeulebroecke, K., Janssens, D., & Swings, J. (2002). Re-examination of the genus *Acetobacter*, with descriptions of *Acetobacter cerevisiae* sp. nov. and *Acetobacter malorum* sp. nov. *Int. J. Syst. Evol. Microbiol. 52*: 1551-1558.

Collins, R. F., Beis, K., Dong, C., Botting, C. H., McDonell, C., Ford, R. C., Clarke, B. R., Whitfield, C., & Naismith, J. H. (2007). The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli. Proc. Natl. Acad. Sci. USA, 104*: 2390-2395.

De Vuyst, L., Lefeber, T., Papalexandratou, Z., & Camu, N. (2008, November). The role of acetic acid bacteria in traditional fermentations. Proceedings of the second international conference on acetic acid bacteria. pp: 24.

Deeraksa, A., Moonmangmee, S., Toyama, H., Adachi, O., & Matsushita, K. (2006).

Conversion of capsular polysaccharide involved in pellicle formation, to extracellular polysaccharide by *galE* deletion in *Acetobacter tropicalis*. *Biosci. Biotechnol. Biochem. 70*: 2536-2539.

Deeraksa, A., Moonmangmee, S., Toyama, H., Yamada, M., Adachi, O., & Matsushita. K. (2005). Characterization and spontaneous mutation of a novel gene, *polE*, involved in pellicle formation in *Acetobacter tropicalis* SKU1100. *Microbiology*, *151*: 4111-4120.

DeLano, W. L. (2002). The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, USA.

Dubois, M., Gilles, K. A., Hamilton, J. K., Robers, P. A., & Smith, F. (1956). Colorimetric method for determination of sugars and related substances. *Anal. Chem. 28*: 350-356.

Duus, J. Ø., Gotfredsen, C. H., & Bock, K. (2000). Carbohydrate structural determination by NMR spectroscopy: Modern methods and limitations. *Chem. Rev. 100*: 4589-4614.

Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M., Pieper, U., & Sali, A. (2006). Comparative Protein Structure Modeling With MODELLER. Current Protocols in Bioinformatics, John Wiley & Sons, Inc. (Suppl. 15, pp. 5.6.1-5.6.30).

Ferreira, A. S., Leitão, J. H., Silva, I. N., Pinheiro, P. F., Sousa, S. A., Ramos, C. G., & Moreira, L. M. (2010). Distribution of cepacian biosynthesis genes among environmental and clinical *Burkholderia* strains and role of cepacian exopolysaccharide in resistance to stress conditions. *Appl. Environ. Microbiol. 76*: 441-450.

Fiser, A., Do, R. K., & Sali, A. (2000). Modeling of loops in protein structures. *Protein Sci. 9*: 1753-1773.

Fry, B. N., Feng, S., Chen, Y. Y., Newell, D. G., Coloe, P. J., & Korolik, V. (2000). The galE gene of *Campylobacter jejuni* is involved in lipopolysaccharide biosynthesis and virulence. *Infect. Immun. 68*: 2594-2601.

Gardy, J. L., Laird, M. R., Chen, F., Rey, S., Walsh, C. J., Ester, M., & Brinkman, F. S. L. (2005). PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics, 21*: 617-623.

Gardy, J. L., Spencer, C., Wang, K., Ester, M., Tusnády, G. E., Simon, I., Hua, S., deFays, K., Lambert, C., Nakai, K., & Brinkman, F. S. L. (2003). PSORT-B: Improving protein subcellular localization prediction for Gram-negative bacteria. *Nucleic Acids Res. 31*: 3613-3617.

Gargiulo, V., De castro, C., Lanzetta, R., Jiang, Y., Xu, L., Jiang, C. L., Molinaro, A., & Parrilli, M. (2008). Structural elucidation of the capsular polysaccharide isolated from *Kaistella flava. Carbohydr. Res. 343*: 2401-2405.

Greenberg, D. E., Porcella, S. F., Stock, F., Wong, A., Conville, P. S., Murray, P. R., Holland, S. M., & Zelazny, A. M. (2006). *Granulibacter bethesdensis* gen. nov., sp. nov., a distinctive pathogenic acetic acid bacterium in the family Acetobacteraceae. *Int. J. Syst. Evol. Microbiol. 56*: 2609-2616.

Greenberg, D. E., Porcella, S. F., Zelazny, A. M., Stock, F., Wong, A., Anderson, V. L., Miller, G., Kleiner, D. E., Tenorjo, A. R., Brinster, L., Dorward., D. W., Murray, P. R., & Holland, S. M. (2006). A novel bacterium associated with lymphadenitis in a patient with chronic granulomatous disease. *PLOS pathog* 2 p e 28.

Grimmecke, H. D., Mamat, U., Lauk, W., Shashkov, A. S., Knirel, Y. A., Vinogradov, E. V., & Kochetkov, N. K. (1991). Structure of the capsular polysaccharide and the O-side-chain of the lipopolysaccharide from *Acetobacter methanolicus* MB58/4 (IMET 10945), and of oligosaccharides resulting from their degradation by the bacteriophage *Acm1. Carbohydr. Res. 220*: 165-172.

Guan, S., Clark, A. J., & Whitfield, C. (2001). Functional analysis of the galactosyltransfreases required for biosynthesis of D-galactan 1, a component of the lipopolysaccharide O1 antigen of *Klebsiella pneumoniae. J. Bacteriol. 183*: 3318-3327.

Hounsell, E. F. (1995). ^{1}H NMR in the structural and conformational analysis of oligosaccharides and glycoconjugates. *Prog. Nucl. Magn. Reson. Spectrosc. 27*: 445-474.

Ishida, T., Sugano, Y., & Shoda, M. (2002). Novel glycosyltransferase genes involved in the acetan biosynthesis of *Acetobacter xylinum. Biochem. Biophys. Res. Commun. 295*: 230-235.

Jennings, M. P., Van der Ley, P., Wilks, K. E., Maskell, D. J., Poolman, J. T., & Moxon, E. R. (1993). Cloning and Molecular analysis of the galE gene of *Nesseria meningitides* and its role in lipopolysaccharide biosynthesis. *Mol. Microbiol. 10*: 361-369.

Jojima, Y., Mihara, Y., Suzuki, S., Yokozeki, K., Yamanaka, S., & Fudou, R. (2004). *Saccharibacter floricola* gen. nov., sp. nov., a novel osmophilic acetic acid bacterium isolated from pollen. *Int. J. Syst. Evol. Microbiol. 54*: 2263-2267.

Kanchanarach, W., Theeragool, G., Inoue, T., Yakushi, T., Adachi, O., & Matsushita, K. (2010). Acetic acid fermentation of *Acetobacter pasteurianus*: Relationship between

acetic acid resistance and pellicle polysaccharide formation. *Biosci. Biotechnol. Biochem. 74*: 1591-1597.

Kanchanarach, W., Theeragool, G., Yakushi, T., Toyama, H., Adachi, O., & Matsushita, K. (2010). Characterization of thermotolerant Acetobacter pasteurianus strains and their quinoprotein alcohol dehydrogenases. *Appl. Microbiol. Biotechnol. 85*: 741-751.

Karatan, E., & Wantick, P. (2009). Signals, regulatory networks, and materials that build and break bacterial biofilms. *Microbiol. Mol. Biol. Rev. 73*: 310-347.

Köplin, R., Brisson, J. R., & Whitfield, C. (1997). UDP-galactopyranose precursor required for formation of the lipopolysaccharide O antigen of *Klebseilla pneumoniae* serotype O1 is synthesized by the product of the rfbD (KP01) gene. *J. Biol. Chem. 272*: 4121-4128.

Kounatidis, I., Crotti, E., Sapountzis, P., Sacchi, L., Rizzi, A., Chouaia, B., Bandi, C., Alma, A., Daffonchio, D., Mavragani-Tsipidou, P., & Bourtzis, K. (2009). *Acetobacter tropicalis* is a major symbiont of olive fruit fly (*Bactrocera oleae). Appl. Environ. Microbiol. 75*: 3281-3288.

Landersjö, C., Yang, Z., Huttunen, E., & Widmalm, G. (2002). Structural studies of the exopolysaccharide produced by *Lactobacillus rhamnosus* strain GG (ATCC 53103). *Biomacromolecules, 3*: 880-884.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Lopez, A. W. R., Thompson, J. D., Gibson T. J., & Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics, 23*: 2947-2948.

Laws, A., Gu, Y., & Marshall, V., (2001). Biosynthesis, characterization, and design of bacterial exopolysaccharides from lactic acid bacteria. *Biotechnol. Adv. 19*: 597-625.

Liu, J., Yu, H., & Shen, Z. (2008). Insights into thermal stability of thermophilic nitrile hydratases by molecular dynamics simulation. *J. Mol. Graph. Model. 27*: 529-535.

Liu, Y. G., & Whittier, R. F. (1995). Thermal asymmetric interlaced PCR: Automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics, 25*: 674-681.

Loganathan, P., & Nair, S. (2004). *Swaminathania salitolerans* gen. nov., sp. nov., a salt-tolerant, nitrogen-fixing and phosphate-solubilizing bacterium from wild rice (Porteresia coarctata Tateoka). *Int. J. Syst. Evol. Microbiol. 54*: 1185-1190.

Manaia, C. M., & Moore, E .R. (2002). *Pseudomonas thermotolerans* sp. nov., a thermotolerant species of the genus *Pseudomonas sensu stricto*. *Int. J. Syst. Evol. Microbiol. 52*: 2203-2209.

Mao, Y., Doyle, M. P., & Chen, J. (2001). Insertion mutagenesis of *wca* reduces acid and heat tolerance of enterohemorrhagic *Escherichia coli* 0157:H7. *J. Bacteriol. 183*: 3811-3815.

Martí-Renom, M. A., Stuart, A .C., Fiser, A., Sánchez, R., Melo, F., & Šali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct. 29*: 291-325.

Matsushita, K., & Ameyama, M. (1982). D-glucose dehydrogenase from *Pseudomonas flucorescens*, membrane-bound. *Methods Enzymol. 89*: 149-154.

Matsushita, K., Ebisuya, H., Ameyama, M., & Adachi, O., (1992). Change of the terminal oxidase from cytochrome a_1 in shaking cultures to cytochrome *o* in static cultures of *Acetobacter aceti*. *J. Bacteriol. 174*: 122-129.

Matsushita, K., Toyama, H., & Adachi, O. (2002). Quinoproteins: Structrue, function, and biotechnological applications. *Appl. Microbiol. Biotechnol. 58*: 13-22.

Metzeger, M., Bellemann, P., Bugert, P., & Geider, K. (1994). Genetic of galactose metabolism of *Erwinia amylovora* and its influence on polysaccharide synthesis and virulence of the fire blight pathogen. *J. Bacteriol. 176*: 450-459.

Moonmangmee, S., Kawabata, K., Tanaka, S., Toyama, H., Adachi, O., & Matsushita, K. (2002a). A novel polysaccharide involved in the pellicle formation of *Acetobacter aceti*. *J. Biosci. Bioeng. 93*: 192-200.

Moonmangmee, S., Toyama, H., Adachi, O., Theeragool, G., Lotong, N., & Matsushita, K. (2002b). Purification and characterization of a novel polysaccharide involved in the pellicle produced by a thermotolerant *Acetobacter* strain. *Biosci. Biotechnol. Biochem. 66*: 777-783.

Nakai, T., Nishiyama, Y., Kuga, S., Sugano, Y., & Shoda, M. (2002). ORF2 gene involves in the construction of high-order structure of bacterial cellulose. *Biochem. Biophys. Res. Commun. 295*: 458-462.

Nassau, P. M., Martin S. L., Brown, R. E., Weston, A., Monsey, D., McNeil, M. R., & Duncan, K. (1996). Galactofuranose biosynthesis in *Escherichia coli* K-12: Identification and cloning of UDP-galatopyranose mutase. *J. Bacteriol. 178*: 1047-1052.

Ndoye, B., Lebecque, S., Dubois-Dauphin, R., Tounkara, L., Guiro, A-T., Kere, C., Diawara, B., & Thonart, P. (2006). Thermoresistant properties of acetic acids bacteria isolated from tropical products of Sub-Saharan Africa and destined to industrial vinegar. *Enzyme Microbiol. Tech. 39*: 916-923.

Niou, Y. K., Wu, W. L., Lin, L. C., Yu, M. S., Shu, H. Y., Yang, H. H., & Lin, G .H. (2009). Role of *galE* on biofilm formation by *Thermus* Spp. *Biochem. Biophys. Res. Commun. 390*: 313-318.

Nishio, Y., Nakamura, Y., Kawarabayasi, Y., Usuda, Y., Kimura, E., Sugimoto, S., Matsui, K., Yamagishi, A., Kikuchi, H., Ikeo, K., & Gojobori, T. (2003). Comparative complete genome sequence analysis of the amino acid replacements responsible for the thermostability of *Corynebacterium efficiens*. *Genome Res. 13*: 1572-1579.

Ohkuri, T., & Yamagishi, A. (2003). Increased thermal stability against irreversible inactivation of 3-isopropylmalate dehydrogenase induced by decreased van der Waals volume at the subunit interface. *Protein Eng. 16*: 615-621

Ophir, T., & Guntnick, D. L., (1994). A role for exopolysaccharides in the protection of microorganisms from desiccation. *Appl. Environ. Microbiol. 60*: 740-745.

Pace, C. N. (1992). Contribution of the hydrophobic effect to globular protein stability. *J. Mol. Biol. 226*: 29-35.

Pan, F., Jackson, M., Ma, Y. F., & McNeil, M. (2001). Cell wall core galactofuran synthesis is essential for growth of mycobacteria. *J. Bacteriol. 183*: 3991-3998.

Pedersen, L. L., & Turco, S. J. (2003). Galactofuranose metabolism: A potential target for antimicrobial chemotherapy. *Cell. Mol. Life Sci. 60*: 259-266.

Perepelov, A. V., Shevelev, S. D., Liu, B., Senchenkova, S. N., Shashkov, A. S., Feng, L., Knirel, Y. A., & Wang, L. (2010). Structures of the O-antigens of *Escherichia coli* O13, O129, and O135 related to the O-antigens of *Shigella flexneri*. *Carbohydr. Res. 345*: 1594-1599.

Poulin, M. B., & Lowary, T. L. (2010). Methods to study the biosynthesis of bacterial furanosides. *Methods Enzymol. 478*: 389-411.

Poulin, M. B., Nothaft, H., Hug, I., Feldman, M., Szymanski, M., and Lowary, T. L. (2009). Characterization of a bifunctional pyranose-furanose mutase from *Campylobacter jejuni* 11168. *J. Biol. Chem. 285*: 493-501.

Raspor, P., & Goranovič, D. (2008). Biotechnological applications of acetic acid bactria. *Crit. Rev. Biotechnol. 28*: 101-124.

Roberts, I. S. (1996). The biochemistry and genetics of capsular polysaccharide production in

bacteria. *Ann. Rev. Microbiol. 50*: 285-315.

Roslund, M. U.,Tähtinen, P., Niemitz, M., & Sjöholm, R., (2008). Complete assignments of the ^{1}H and ^{13}C chemical shifts and $J_{H,H}$ coupling constants in NMR spectra of D-glucopyranose and all D-glucopyranosyl-D-glucopyranosides. *Carbohydr. Res. 343*: 101-112.

Saeki, A., Theeragool, G., Matsushita, K., Toyama, H., & Adachi, O. (1997). Development of thermotolerant acetic acid bacteria useful for vinegar fermentation at higher temperatures. *Biosci. Biotechnol. Biochem. 61*: 138-145.

Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol. 234*: 779-815.

Sandford, P. A., & Baird J. (1983). Industrial utilization of polysaccharides in The Polysaccharides (Aspinall, G.O., ed) pp: 411-490.

Sassaki, G. L., Iacomini, M., & Gorin, P. A. J. (2005). Methylation-GC-MS analysis of arabinofuranose-and galactofuranose-containing structures: rapid synthesis of partially *O*-methylated alditol acetate standards. *Annals of the Brazilian Academy of Sciences*, *77*: 223-234.

Saxena, I. M., Kudlicka, K., Okuda, K., & Brown, R. M., (1994). Characterization of genes in the cellulose-synthesizing operon (*acs* operon) of *Acetobacter xylinum*: Implications for cellulose crystallization. *J. Bacteriol. 176*: 5735-5752.

Sikorski, J., Brambilla, E., Kroppenstedt, R. M. & Tindall, B. J. (2008). The temperature-adaptive fatty acid content in *Bacillus simplex* strains from 'Evolution Canyon', Israel. *Microbiology 154*: 2416-2426.

Sutherland, I. W. (1994). Structure-function relationships in microbial exopolysaccharides. *Biotechnol. Adv., 12*: 393-448.

Sutherland, I.W. (1998). Novel and established applications of microbial polysaccharides. *Trends Biotechnol. 16*: 41-46.

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

Weston, A., Stern, R. J., Lee, R. E., Nassau, P. M., Monsey, D., Martin, S. L., Scherman, M. S., Besra, G. S., Duncan, K., & McNeil, M. R. (1998). Biosynthetic origin of mycobacterial cell wall galactofuranosyl residues. *Tuber. Lung Dis. 78*: 123-131.

Wong, H. C., Fear, A. L., Calhoon, R. D., Eichinger, G. H., Mayer, R., Amikam, D., Benzima, M., Gelfand, D. H., Meade, J. H., Emerick, A. W., Bruner, R., Ben-Bassat, A., & Tal, R. (1990). Genetic organization of the cellulose synthase operon in

Acetobacter xylinum. *Proc. Natl. Acad. Sci. USA*, *87*: 8130-8134.

Wozniak, D. J., Wyckoff, T. J. O., Starkey, M., Keyser, R., Azadi, P., O'Toole, G. A., & Parsek, M. R. (2003). Alginate is not a significant component of the extracellular polysaccharide matrix of PA14 and PA01 *Pseudomonus aeruginosa* biofilms. *Proc. Natl. Acad. Sci. USA*, *100*: 7907-7912.

Xayarath, B., & Yother, J. (2007). Mutations blocking side chain assymbly, polymerization, or transport of a Wzy-dependent *Streptococcus pneumoniae* capsule are lethal in the absence of suppressor mutations and can affect polymer transfer to the cell wall. *J. Bacteriol. 189*: 3369-3381.

Yamada, Y., & Yukphan, P. (2008). Genera and species in acetic acid bacteria. *Int. J. Food Microbiol. 125*: 15-24.

York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T., & Albersheim, P. (1985). Isolation and characterization of plant cell walls and cell wall components. *Methods Enzymol. 118*: 3-40.

Yoshida, T., Ayabe, Y., Yasunaga, M., Usami, Y., Habe, H., Nojiri, H., & Omori, T. (2003). Genes involves in the synthesis of the exopolysaccharide methanolan by the obligate methylotroph *Methylobacillus* sp. strain 12S. *Microbiology*, *149*: 431-444.

Yukphan, P., Malimas, T., Muramatsu, Y., Takahashi, M., Kaneyasu, M., & Potacharoen, W. (2009). *Ameyamaea chiangmaiensis* gen. nov., an acetic acid bacterium in the α-proteobacteria. *Biosci. Biotechnol. Biochem. 73*: 2156-2162.

Yukphan, P., Malimas, T., Muramatsu, Y., Takahashi, M., Kaneyasu, M., & Tanasupawat, S. (2008). *Tanticharoenia sakaeratensis* gen. nov, sp. Nov., a new osmotolerant acetic acid bacterium in the α-proteobacteria. *Biosci. Biotechnol. Biochem. 72*: 672-676.

Yukphan, P., Malimas, T., Potacharoen, W., Tanasupawat, S., Tanticharoen, M., & Yamada, Y. (2005). *Neoasaia chiangmaiensis* gen. nov., sp. nov., a novel osmotolerant acetic acid bacterium in the alpha-Proteobacteria. *J. Gen. Appl. Microbiol. 51*: 301-311.

ACKNOWLEGEMENTS

I would like to express my deepest gratitude to my supervisor professor Kazunobu Matsushita for having me in his laboratory and for his guidance throughout my study. I am also grateful to Professor Mamoru Yamada for accepting me to study in this program and for his help and advice. I would very much like to thank Emeritus Professor Osao Adachi for his encouragement and support during my course of study.

Thanks are due to Associated Professor Yoshihiko Akakabe from the department of biological chemistry for helping with NMR analysis. I would also like to acknowledge Associated Professor Toshiharu Yakushi and Dr. Tomoyuki Kosaka for their useful discussion and suggestions. I am also grateful to Dr. Minenosuke Matsutani for helping in bioinformatic analysis. I am very indebted to my committee members for their time and useful suggestions.

I wish to express my thanks to Monbukagakusho (Japan Government Scholarship for Foreign Students) for financial support and to Industrial Research and Consultancy Center for giving me a temporal leave to study in this program.

I would like to thank Dr. Somporn Moonmangmee, Dr. Arpaporn Deeraksa, Dr. Wichai Soemphol, and Dr. Natsaran Saichana for useful discussion and help. My thanks are extended to all former and present members of obi laboratory (Laboratory of Applied Microbiology) and jose laboratory for joint seminars, sports, and help. I also want to appreciate my friends in Yamaguchi for being there for me.

Last but not least I would like to thank my family for unconditional love and support throughout my life.

SUMMARY

Structure and biosynthesis of pellicle polysaccharides in a thermotolerant acetic acid bacterium, *Acetobacter tropicalis* SKU1100

Acetobacter tropicalis SKU1100 is a thermotolerant acetic acid bacterial strain capable of producing pellicle even at higher temperature. The pellicle is an air-liquid interface biofilm that contains bacterial cells attached to the surface and to each other by polysaccharides and other extracellular matrix. Capsular polysaccharide (CPS) was produced for pellicle formation in *A. tropicalis* SKU1100. Genetic study revealed that *polABCDE* operon is involved in CPS biosynthesis. The *polABCD* showed high similarity to *rfbBACD* operon which is responsible for d-TDP-rhamnose biosynthesis, while the *polE* is a novel gene exhibiting a low similarity to other glycosyltransferase genes. Disruption of *polE* abolished pellicle formation and CPS production, instead the mutant strain produced EPS that secreted in the growth medium. Likewise, disruption of *galE* gene, encoding UDP-galactose-4-epimerase, led to EPS production and no pellicle formation. This study was implemented to gain insight into thermotolerance and pellicle polysaccharide biosynthesis of *A. tropicalis* SKU1100.

In Chapter 1, to understand the mechanism of thermotolerance in *A. tropicalis*, the draft genome of the strain was obtained, and compared with the genome of a mesophilic strain, *Acetobacter pasteurianus* IFO3283-01. The comparative genomic study showed that amino acid substitutions from large to small residues and lysine to arginine occur in many orthologous genes in *A. tropicalis* SKU1100. Moreover, comparative modeling study between orthologous proteins indicated that the number of arginine-based salt bridges increased in proteins of *A. tropicalis* compared with *A. pasteurianus*. Since it has been reported that arginine-based salt bridges are important factor for thermo-stability of protein structure, the results suggest that the increased number of arginine-based salt bridges may contribute to the thermotolerance of *A. tropicalis* SKU1100 most probably by stabilizing protein structure.

In Chapter 2, to understand PolE function and CPS biosynthesis, the structures of wild type CPS and Δ*polE* mutant EPS were examined using monosaccharide composition analysis, methylation analysis, and ^{1}H and ^{13}C NMR spectroscopy. The methylation analysis of the purified polysaccharides indicated that wild type CPS contains 2,3-linked rhamnose residue, terminal and 2-linked glucose residues, terminal furanose residue, and 6-linked

galactopyranose residue as major components, and some minor components, double branched 2,3,4-rhamnose, terminal galactopyranose, and 3-linked rhamnose. EPS of Δ*polE* mutant had the same major components as those of the CPS and increased number of minor components such as glucose, galactose, and rhamnose residues. Thus, wild type CPS and Δ*polE* mutant EPS were shown to be composed of the same major residues of 2,3-rhamnose, 2-glucose, terminal-glucose, terminal galactofuranose (Galf), and 6-galactose. Furthermore, NMR analysis suggested that the two polysaccharides have a branched hexasaccharide repeating unit composed of 2,3-α-L-rhamnopyranose (2 moles), 2-α-D-glucopyranose, 6-β-D-galactopyranose, and two branched residues of terminal-β-D-Galf and terminal-α-D-glucopyranose. By comparing these structures, one 2,3,4-α-L- rhamnopyranose double branched residue observed in wild type CPS was shown missing in the mutant EPS. Thus PolE may function as β-galactosyltransferase that adds β- galactopyranose to 2,3-α-L-rhamnopyranose to make the double branched 2,3,4- α-L-rhamnosyl residue, which leads to CPS formation.

In Chapter 3, the structure of Δ*galE* mutant EPS was studied, as newly produced EPS applicable in industry. Methylation analysis indicated that Δ*galE* mutant EPS contains 2,3-linked rhamnose, 3-linked rhamnose, terminal and 2-linked glucose residues as the major peaks and with some minor peaks of glucose, and rhamnose. The absence of the furanose residue was observed in Δ*galE* mutant EPS, supporting the notion that the terminal furanose residue observed in CPS and Δ*polE* EPS is a galactofuranose as Δ*galE* mutant cannot produce galactose. Thus, the EPS was composed of 2,3-rhamnose, 2-glucose, terminal glucose, and 3-rhamnose as major components plus minor glucose and rhamnose residues. NMR analysis suggested it has a branched tetrasaccharide repeating unit composed of 2,3-α-L-rhamnopyranose, 2-α-D-glucopyranose, and 3-α-L-rhamnopyranose with terminal- α-D-glucopyranose branch. Hence, this EPS seems to be derived from the wild type CPS by missing the galactose residues.

In Chapter 4, I investigated Galf biosynthesis by searching *glf* gene that encodes UDP-galactopyranose mutase responsible for Galf biosynthesis from galactopyranose. The part of *glf* gene was cloned by PCR using primers designed from the conserved regions of *glf* from many bacteria, and the whole length gene by using TAIL PCR method. To study the function of this gene, deletion mutant was prepared using fusion PCR method. However, the mutant strain, Δ*glf*, could still produce pellicle and CPS. Moreover, NMR analysis confirmed the presence of Galf residue in the CPS of the mutant. These results together with the homology search in the draft genome sequence of *A. tropicalis* SKU1100 suggest the

existence of another paralogous gene that may work with the identified gene or individually to produce Galf. Additional study is needed to identify the paralog gene and study its function.

和文要旨

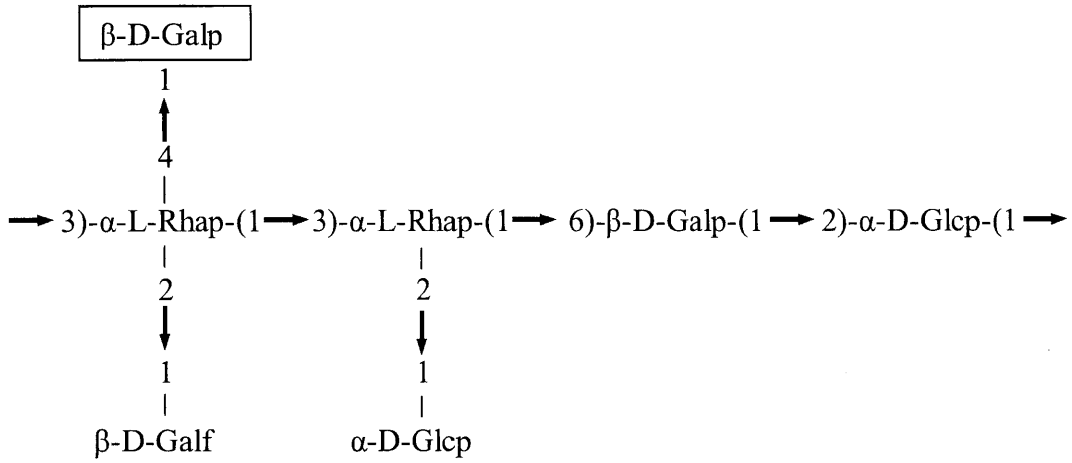
Structure and biosynthesis of pellicle polysaccharides in a thermotolerant acetic acid bacterium, *Acetobacter tropicalis* SKU1100

Acetobacter tropicalis SKU1100は高温下で高い菌膜生成能を有する耐熱性酢酸菌である。菌膜は静置培養表面に形成されるバイオフィルムであり、菌が生成する多糖やその他の細胞外排出物と菌体によって形成されている。*A. tropicalis* SKU1100は細胞に結合した多糖（Capsular polysaccharide: CPS）より成る菌膜を形成する。本菌の遺伝子研究より、*polABCDE* オペロンがその CPS 形成に関与していることが示されている。このオペロンのうち、*polABCD* は d-TDP-rhamnose の生合成に関与することが知られている *rfbBACD* オペロンに高い相同性を有しているが、*polE* は glycosyltransferase 遺伝子に低い相同性を示している。この *polE* 欠損株は、本菌の CPS 生成能を失わせ、菌膜生成能を消失させる。しかし、新たに菌体外多糖（Extracellular polysaccharide: EPS）を生成するように変化する。また、UPD-galactose 4-epimerase をコードする *galE* 遺伝子の欠損も菌膜生成能を失わせ、EPS を生成するようになることが知られている。そこで、本研究は、本菌の耐熱性と菌膜多糖生合成についての知見を得るために企画された。

第１章では、本菌の耐熱機構を理解するために、本菌のドラフトゲノム解析を行い、中温菌 *Acetobacter pasteurianus* IFO3283-01 株ゲノムとの比較を行った。その結果、両株に共通する多くのオルソログ遺伝子において、本菌の遺伝子で大きなアミノ酸から小さな残基へ、リジンからアルギニンへの置換が多く蓄積していることがわかった。さらに、これらのオルソログ・タンパク質のいくつかについて、モデル構造解析を行うことで、本菌のタンパク質でアルギニンによる塩架橋が増加していることがわかった。アルギニン塩架橋は、タンパク質の熱安定化に必要なことが知られているので、この塩架橋の増加が本菌の耐熱性に貢献していると予想された。

第２章では、本菌の PolE タンパク質の機能および CPS 生合成系を理解するために、野生株の CPS および *polE* 欠損株の EPS の構造解析が行われた。両多糖とも、以下に示すような、2 分子の 2,3-α-L-Rhamnopyranose, 各 1 分子の 2-α-D-

Glucopyranose, 6-β-D-Galactopyranose, Terminal-β-D-Galactofuranose, Terminal-α-D-Glucopyranose を含む６糖の繰り返しユニットから成ることを、NMR 解析を含むいくつかの手法を組み合わせることで明らかにした。また、野生株 CPS と変異株 EPS の両構造を比較することよって、CPS には 2,3,4-α-L-Rhamnopyranose による２重分枝鎖が存在することが示され、PolE タンパク質は β-Galactopyranose（枠で囲まれた糖）を 2,3-α-L-Rhamnopyranose 残基に結合させる β-galactosyltransferase ではないかと推測された。この結合によって、CPS は細胞に結合していると推定された。



第３章では、同様な手法を用いて、*galE* 欠損株の EPS 構造解析が行われ、2,3-α-L-Rhamnopyranose, 2-α-D-Glucopyranose, 3-α-L-Rhamnopyranose, Terminal-α-D-Glucopyranose からなる４糖繰り返しユニットから成ることがわかった。このことは、本変異株では、UPD-Galactose 分子が生成できないため、ガラクトースの欠失した多糖となり、かつ細胞結合能を失って EPS になると考えられた。

第４章では、Galactofuranose (Galf) 合成に関与すると考えられる UDP-galacto-pyranose mutase をコードする *glf* 遺伝子のクローニングが行われた。PCR および TAIL-PCR を組み合わせて、その全長をクローニングし、その遺伝子を使って、*glf* 欠損株を作成した。しかしながら、その欠損株は菌膜を作る能力を維持しており、その CPS の NMR 解析から Galf が存在していることも確認された。本菌株ゲノムには *glf* のパラログの存在が確認されたため、今後、それらの解析が引き続き必要であることがわかった。

LIST of PUBLICATIONS

1- **Ibnaof Ali Ibnaof Ali**, Yoshihiko Akakabea, Somporn Moonmangmee, Arpaporn Deeraksaa, Minenosuke Matsutania, Toshiharu Yakushia, Mamoru Yamada, and Kazunobu Matsushita: Structural characterization of pellicle polysaccharides of *Acetobacter tropicalis* SKU1100 wild type and mutant strains.

Carbohydrate Polymers 86 (2011) 1000-1006.

(Chapter 2 and 3)

2- Minenosuke Matsutani, Hideki Hirakawa, Mitsuteru Nishikura, Wichai Soemphol, **Ibnaof Ali Ibnaof Ali**, ToshiharuYakushi, and Kazunobu Matsushita: Increased number of Arginine-based salt bridges contributes to the thermotolerance of thermotolerant acetic acid bacteria, *Acetobacter tropicalis* SKU1100

Biochemical and Biophysical Research Communications 409 (2011) 120-124.

(Chapter 1)