HOKKAIDO UNIVERSITY

Title	Studies on alkane degradation and its molecular mechanisms of Geobacillus kaustophilus HTA426 isolated from the Mariana Trench
Author(s)	NITHIMETHACHOKE, Tanasap
Degree Grantor	北海道大学
Degree Name	博士(環境科学)
Dissertation Number	甲第16030号
Issue Date	2024-06-28
DOI	https://doi.org/10.14943/doctoral.k16030
Doc URL	https://hdl.handle.net/2115/95481
Type	doctoral thesis
File Information	NITHIMETHACHOKE_Tanasa.pdf



Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Studies on alkane degradation and its molecular mechanisms of *Geobacillus kaustophilus* HTA426 isolated from the Mariana Trench

（マリアナ海溝から単離された *Geobacillus kaustophilus* HTA426 のアルカン分解とその分子機構に関する研究）

Doctoral dissertation

Tanasap Nithimethachoke

Division of Biosphere Science
Graduate School of Environmental Science
Hokkaido University, Japan

2024

ABSTRACT

The genus *Geobacillus* is a Gram positive thermophilic spore-forming bacterial group, which has gained attention due to its high-temperature tolerant nature and its benefits in the industrial production of thermostable enzymes. Several *Geobacillus* strains have been reported to degrade hydrocarbon. In a previous study on *Geobacillus thermoleovorans* B23 (Boonmak et al., 2014), a broad range alkane (C11 – C32) degrader, it was suggested that *Geobacillus kaustophilus* HTA426 can also degrade alkane. However, genome analysis of *G. kaustophilus* HTA426 suggested that no known alkane monooxygenase enzymes exist in the genome, e.g. AlkB-type, Alm-type, CYP-type, and Lad-type which can be found in other alkane degrading Geobacilli. I identified a candidate gene, *GK2771*, in the HTA426 genome that encodes an isolated ribonucleotide reductase small subunit (RNR2) located in a cluster of genes responsible for alkane degradation. Recently, GK2771 has been proposed as a member of a new protein group, R2-like ligand binding oxidase (R2lox), which resembles RNR2 class Ic but lacks ribonucleotide reductase activity. However, the function and substrate for R2lox are yet unknown. I hypothesized that it could function as an alkane monooxygenase, based on its structural high similarity to soluble methane monooxygenase (sMMO).

Chapter 1, it was verified that *G. kaustophilus* HTA426 can degrade alkane (C11 – C17) in nutrient medium LB at 60℃. Moreover, it was found to degrade wider range C10 – C24 alkane in minimal medium LBM condition.

Chapter 2, recombinant *E. coli* (pET28a-GK2771) was prepared in which *GK2771* is overexpressed under the control of T7-promoter. As expected, the recombinant GK2771 crude enzyme degraded alkane and its activity was not much different between 37℃ and 60℃. This finding demonstrates that an isolated RNR2, GK2771 (*Gk*R2loxI), is a novel alkane monooxygenase, and answers questions about the function and substrate of the R2lox protein group.

Chapter 3, *gk2772*, a putative aldehyde dehydrogenase gene, was overexpressed in *E. coli* (pET28a-GK2772), purified and examined its activity. Recombinant GK2772 showed significantly higher activity against acetaldehyde at 60℃ compared to 30℃, 40℃ and 50℃ conditions. It was also suggested that GK2772 has dehydrogenase activity against fatty aldehyde, decanal C10 and C14.

Chapter 4, describes general discussion including distribution of GK2771 gene homologues. This study expands our understanding of the vast diversity and new evolutionary lineage of the bacterial alkane monooxygenase family.

TABLE OF CONTENTS

PROLOGUE

0.1 Alkane 5

0.2 Alkane degradation in bacteria 6
0.2.1 Anaerobic alkane degradation 6
0.2.2 Aerobic alkane degradation 7

0.3 The Geobacillus 9
0.3.1 *Geobacillus thermoleovorans* B23 9
0.3.1 *Geobacillus kaustophilus* HTA426 10

0.4 The isolated RNR2, GK2771 11

0.5 R2-like ligand-binding oxidases (R2lox) 13

CHAPTER I Alkane degradation by *Geobacillus kaustophilus* HTA426

1.1 Materials and methods
1.1.1 Growth of *G. kaustophilus* HTA426 on minimal medium 16
1.1.2 Alkane degradation test 16
1.1.3 Gas chromatography 17

1.2 Results
1.2.1 Growth of *G. kaustophilus* HTA426 on minimal medium 17
1.2.2 Alkane degradation test 17

1.3 Discussion 20

CHAPTER II Heterologous gene expression of GK2771 in *E. coli* BL21 (DE3)

2.1 Materials and methods
2.1.1 Strains and plasmids 24
2.1.2 Cloning of *gk2771* and construction of expression plasmid
2.1.2.1 *gk2771* 27
2.1.2.2 Mutated *gk2771* 29
2.1.2.3 *gk2771* with His tag 30
2.1.3 Heterologous gene expression in *E. coli.* 31
2.1.4 Protein purification
2.1.4.1 GK2771 and mutated GK2771 31
2.1.4.2 Histidine tagged GK2771 32
2.1.5 Alkane degradation assay 33

2.2 Results
2.2.1 Cloning of GK2771 gene and construction of expression plasmid 34
2.2.2 Heterologous gene expression of GK2771 and its variations in *E. coli* 36
2.2.3 Protein purification 38
2.2.4 Alkane degradation assay 41

2.3 Discussion 44

CHAPTER III Heterologous gene expression of GK2772 in *E. coli*

3.1 Materials and methods
3.1.1 Strains and plasmids 46
3.1.2 Cloning of GK2772 gene and construction of gene expression plasmid 47
3.1.3 Heterologous gene expression of GK2772 in *E. coli* 48
3.1.4 Protein purification 48
3.1.5 Aldehyde dehydrogenase assay 49

3.2 Results
3.2.1 Cloning of GK2772 gene and construction of gene expression plasmid 50
3.2.2 Heterologous gene expression of GK2772 in *E. coli* 51
3.2.3 Protein purification 52
3.2.4 Aldehyde dehydrogenase assay 53

3.3 Discussion 61

CHAPTER IV General discussion 62

ACKNOWLEDGMENT 69

REFERENCES 70

PROLOGUE

0.1 Alkane

Pollution of the environment by crude oil spillage has long been, and is still today, a concerning problem. Crude oil is the base for significantly wide range of products, mainly fuels such as gasoline and diesel, and daily products like plastics and rubbers. Therefore, crude oil usage has been constantly increasing. During petroleum processing, accidental spills and discharges of petroleum hydrocarbons often occur (Xu et al., 2018), which result in environmental pollution both on land and in the sea. This causes damage to the environment and living organism. To deal with these problems, bioremediation has become a good choice due to its low cost and eco-friendly nature (Guerra et al., 2018). Alkanes are one of the major component groups of crude oil which lead to extensively studying of alkane degrading organisms.

Alkanes are saturated aliphatic hydrocarbons, consisting only of carbon and hydrogen atoms. They can be linear (n-alkanes), cyclic (cyclo-alkanes) or branched (isoalkanes). Alkanes are the most abundant hydrocarbons (20–50%) in crude oil. As alkanes are non-polar molecules with very inert chemical activity, their utilization by microorganisms irrespective of assimilation and dissimilation faces significant challenges. Alkane can be described in three categories according to the length of carbon chain. However, the definition was imprecise and varied from each publication (Wentzel et al., 2007). In this dissertation, the alkane will be categorized as follows: 1. Short chain alkane: C1–C10, 2. Medium chain alkane: C11–C19, and 3. Long chain alkane: from C20 and over.

0.2 Alkane degradation in bacteria

0.2.1 Anaerobic alkane degradation

By bacteria, alkane degradation could occur under both aerobic and anaerobic conditions. In this dissertation the focus will be on aerobic reactions therefore the anaerobic condition will be described here briefly. Under anaerobic conditions, nitrate or sulfate is used as an electron acceptor instead of oxygen (Ji et al., 2013). There are two main strategies for anaerobic alkane activation 1. by addition to fumarate, catalyze by alkylsuccinate synthase (ass) or (1-methylalkyl)succinate synthase (mas), addition of fumarate to n-alkane yields alkyl succinate. 2. via carboxylation, with help of bicarbonate subterminal carboxylation occurs at C3 position. The end product of both reactions is fatty acid which then undergoes Beta-oxidation (Fig.1) (Boll et al., 2002 and Mbadinga et al., 2011).

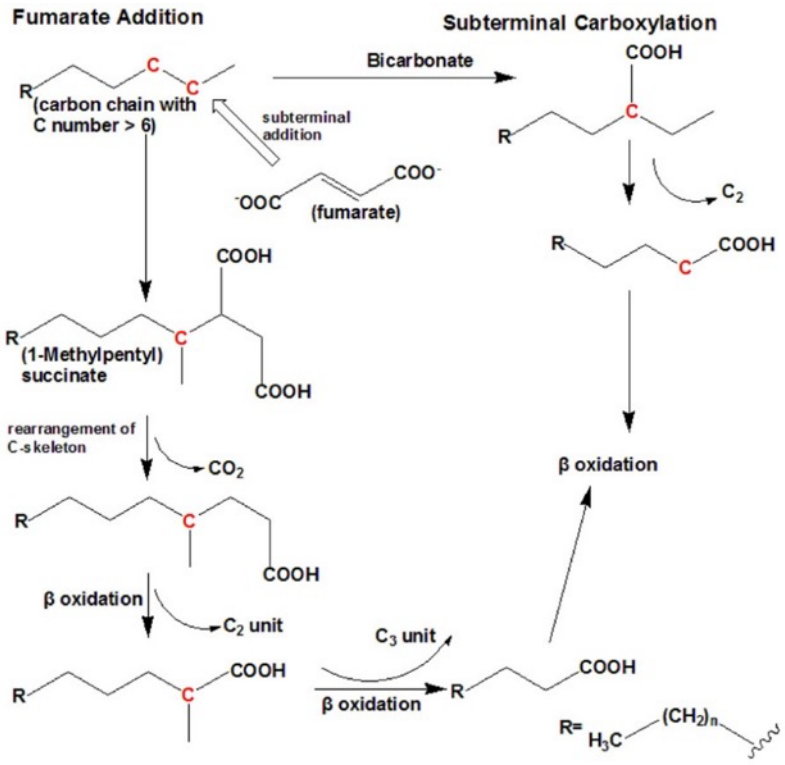


Fig. 1. The anaerobic alkane degradation pathway (Ji et al., 2013)

Two types of anaerobic degradation pathway that currently have been identified. Fumarate addition (left) and Subterminal carboxylation (right). Reprinted with permission from Ji et al., 2013. Copyright 2024 Frontiers Media S.A.

0.2.2 Aerobic alkane degradation

In the presence of oxygen, the reaction is initiated by oxygenase enzymes (Table 1). These enzymes can be divided into groups based on the alkane length, it is possible for enzymes to degrade multiple types of alkanes or even other types of hydrocarbons (van Beilen and Funhoff, 2007; Fig. 2).

First, short chain alkane hydroxylating enzyme, methane monooxygenase (MMO) which has two distinct types; membrane bound particulate methane monooxygenase (pMMO) and soluble methane monooxygenase (sMMO). Even though pMMO can be found in almost methanotrophs, it is less studied than sMMO due to difficulties of isolating stable and active purified samples. sMMO consists of three components: a hydroxylase unit (MMOH), a reductase unit (MMOR), and a regulatory protein (MMOB), it has a non-heme diiron active site. While pMMO also has three components but all is hydroxylase (PmoA, PmoB and PmoC) and uses copper as an active site. Some strains contain both types of MMO, an expression of enzyme is regulated by an availability of the corresponding metal ions (Ross & Rosenzweig, 2017; Semrau, DiSpirito, & Yoon, 2010; Sirajuddin & Rosenzweig, 2015).

Second, medium chain alkane hydroxylases are AlkB and Cytochrome P450. The Alk system consists of integral membrane non-heme diiron monooxygenase (AlkB), electron transfer protein, rubredoxin (AlkG) and rubredoxin reductase (AlkT). The cytochrome P450 alkane hydroxylase belongs to CYP153 family, a soluble heme monooxygenase that requires ferredoxin and ferredoxin reductase. Both gene homologues can be found across many bacteria group, also well characterized (Liang et al.,2016; Rojo, 2009). It is worth to note that P450 monooxygenase is also one of important enzymes in human liver responsible for detoxification and biotransformation of xenobiotics and drugs (https://doi.org/10.1146/annurev.pharmtox.45.120403.100030).

The last group, long chain alkane hydroxylases, are AlmA and LadA types. AlmA is a flavin-binding monooxygenase, the genes activity was confirmed but still poorly understood. LadA is a two component flavin dependent oxygenase which consists of a NAD(P)H dependent flavin reductase and the LadA monooxygenase. There are still many bacteria that can degrade

long chain alkane that gene have not been identified and that could be unrelated to what is currently known (Elumalai et al., 2017; Rojo, 2009).

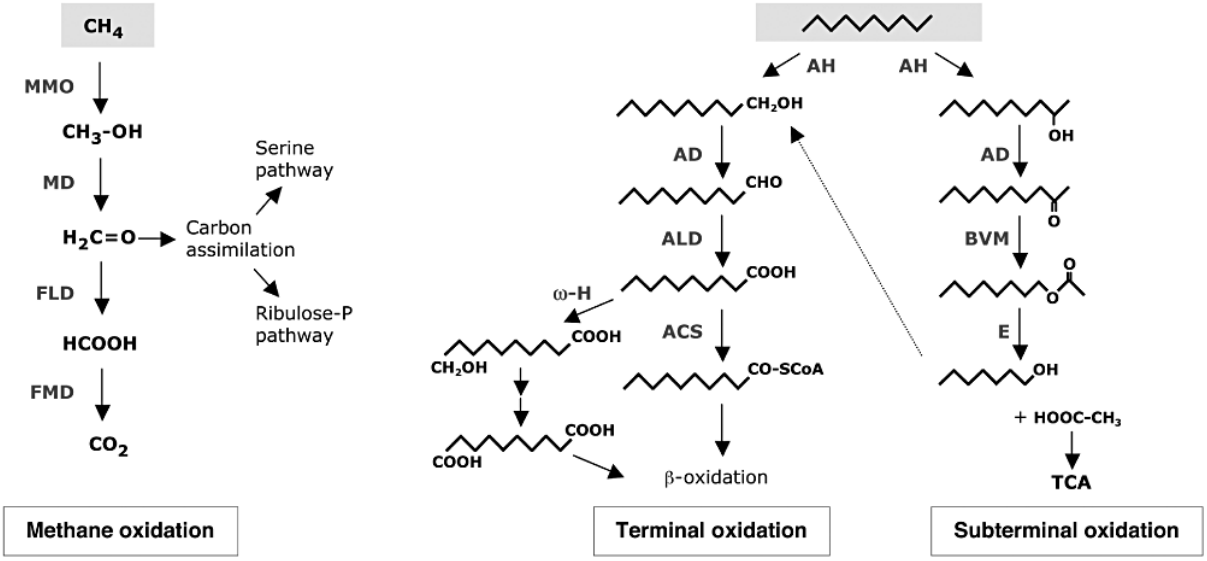


Fig. 2. Aerobic degradation pathways of methane (left), and of larger n-alkanes by terminal and subterminal oxidation (right) Initial activation of the alkane molecule requires O_2 as a reactant. Abbreviation is as follows: MMO, methane monooxygenase; MD, methanol dehydrogenase, FLD, formaldehyde dehydrogenase; FMD, formate dehydrogenase. AH, alkane hydroxylase; AD, alcohol dehydrogenase; ALD, aldehyde dehydrogenase; ACS, acyl-CoA synthetase; w-H, w-hydroxylase; BVM, Baeyer–Villiger monooxygenase; E, esterase; TCA, tricarboxylic acids cycle (Rojo, 2009). Reprinted with permission from Rojo, 2009. Copyright 2024 John Wiley & Sons, Inc.

Table 1 Enzymes involved in the hydroxylation of alkanes. (Chanita Boonmak, 2014)

Enzyme	Composition and cofactors	Substrate range	Presence shown in the genus
Soluble methane monooxygenase (sMMO)	α2β2γ2 hydroxylase; diiron reductase, [2Fe–2S], FAD, NADH regulatory subunit	C1–C8 (halogenated)-alkanes, alkenes, cycloalkanes	*Methylococcus, Methylosinus, Methylocystis, Methylomonas, Methylocella*
Particulate methane monooxygenase (pMMO)	α3β3γ3 hydroxylase trimer composed of PmoA, PmoB, PmoC;	C1–C8 (halogenated)-alkanes, alkenes, cycloalkanes	*Methylococcus, Methylosinus, Methylocystis, Methylobacter, Methylomonas, Methylomicrobium,* etc.
AlkB-type alkane monooxygenase	Membrane monooxygenase; dinuclear iron rubredoxin; rubredoxin reductase, FAD, NADH	C5–C16 alkanes, fatty acids, alkylbenzenes, cycloalkanes, etc.	*Acinetobacter, Alcanivorax, Burkholderia, Geobacillus, Mycobacterium,Pseudomonas, Rhodococcus,* etc.
AlmA-type alkane monooxygenase	Soluble AlmA alkane monooxygenase, NADH/flavin reductase	C10–C40 alkanes	*Acinetobacter*

LadA-type alkane monooxygenase	Soluble LadA alkane monooxygenase, (homo-dimer), $FMNH_2$	C15–C36 alkanes	*Geobacillus thermodenitrificans* NG80-2 and *Geobacillus thermoleovorans* B23 etc.
Bacterial P450 monooxygenase systems (CYP153, class I)	P450 monooxygenase; P450 heme ferredoxin; iron–sulfur ferredoxin reductase, FAD, NADH	C5–C16 alkanes, (cyclo)-alkanes, alkylbenzenes, etc.	*Acinetobacter, Alcanivorax, Amycolicicoccus subflavus, Caulobacter, Mycobacterium, Rhodococcus, Sphingomonas,* etc.

0.3 Genus *Geobacillus*

Geobacillus is a group of thermophilic bacteria, whose growth temperature ranges from 35 to 75 °C and optimum at 55 to 65 °C, Gram-positive spore-forming aerobic or facultative anaerobic. The genus *Geobacillus* is proposed and separated from mesophilic genus *Bacillus* in 2001. Due to thermophilic characteristics and simple nutrient requirement, several *Geobacillus* spp. are used to produce products at high temperature conditions which production are inefficient at room temperature and also sources of many thermostable enzymes. These are amylases, lipases, β-galactosidases, cellulases, xylanases, endonucleases, and proteases (DeLong, 2013). Many *Geobacillus* bacteria were isolated from oilfield environments, this could be a possible reason that these microbes have hydrocarbon uptake and degradation capacity. However, not all *Geobacbillus* species can degrade hydrocarbon. Alkane degradation ability of each *Geobacillus* is various and by different gene, some can degrade medium chain alkane, long chain alkane or even aromatic hydrocarbon (Lin, Zhang, Tao, Wang, & Li, 2019). For example, many strains or closely related of *Geobacillus stearothermophilus* is able to degrade medium to long chain alkane or even aromatic like naphthalene, with different preference and having AlkB type gene for alkane hydroxylase (Liu et al., 2009; Sun et al., 2015; Zhou et al., 2018). *Geobacillus thermodenitrificans* NG80-2 and *Geobacillus thermoleovorans* B23 contain different LadA homolog, degrading C15-C36 and C11-C32 respectively (Boonmak et al., 2014; Feng et al., 2007).

0.3.1 *Geobacillus thermoleovorans* B23

During a study of LadA type alkane monooxygenase genes in *G. thermoleovorans* B23 (Boonmak et al., 2014), it was suggested that *Geobacillus kaustophilus* HTA426 have alkane degradation activity (Fig. 3, Fig. 4). *G. thermoleovorans* B23 is a thermophilic bacterium, isolated from 2,150 m depth of subterranean petroleum reservoir in Nigata, Japan, 2001. It can degrade aliphatic alkanes (C11 - C32), containing multiple LadA type genes, $LadA\alpha_{B23}$,

LadAβ$_{B23}$, and *LadB*. In heterogous expression experiment of *G. thermoleovorans* B23's *LadAα*$_{B23}$, a negative control, non- expression vector only host *G. kaustophilus* MK93 was also able to degrade n-eicosane (C20) at 70˚C in minimal medium. *G. kaustophilus* MK93 is a derivative strain of *G. kaustophilus* HTA426, constructed for counter selection system of transformation. β-amylase gene (GK0707) was inserted by α-galactosidase gene (GenBank accession no. M13466, bgaB) from *G. stearothermophilus* IAM11011 with maltose promoter of HTA426 (Pgk704), *ΔpyrF ΔpyrR GK707::Pgk704 bga*). It was constructed as a model for high temperature microbial processes (HTMPs) (Suzuki et al.,2013). This incident made me question about the parental strain, *G. kaustophilus* HTA426 alkane degradation ability due to no known alkane degradation gene is presented in its chromosome.

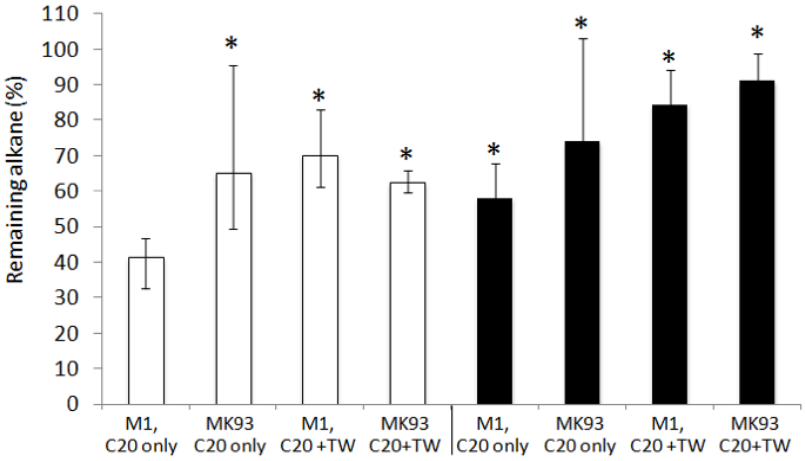


Fig. 3. Remaining alkane after cultivated host *G. kaustophilus* strain MK93 and recombinant strain M1, harboring *frd* and *ladA*α$_{B23}$ genes on Minimal salts medium (Suzuki *et al.,* 2013) supplemented with 1% (vol/vol) of *n*-eicosane (C20 alkane) at 70°C for 3 days. □, medium containing 1% (vol/vol) of maltose; ■, medium containing 1% (vol/vol) of glucose. *$P < 0.05$ compared with the positive recombinant M1 on medium containing 1% (vol/vol) maltose without Tween 80 (Boonmak, 2014). M1 degraded more C20 alkane than MK93, however, MK93 unexpectedly showed alkane degradation activity.

0.3.2 *Geobacillus kaustophilus* HTA426

G. kaustophilus HTA426 is a thermophilic and aerobic endospore-forming Gram-positive *Bacillus*-related species, whose upper temperature limit for growth is 74 °C (optimally at 60 °C in LB), isolated from deep sediment of the Mariana Trench. Its genome is used in comparative genomic studies with other mesophilic bacilli to facilitate the study of the thermophilicity and thermostability of proteins, because this was the first genomic sequence

from a thermophilic Bacillus-related species (Takami et al.,2004). After searching through the HTA426 genome, no conventional alkane hydroxylase genes were found. Based on what we observed from *G. kaustophilus* MK93, preliminary experiment on alkane degradation of *G. kaustophilus* HTA426 had been done, in LB media for 5 and 10 days at 60℃, compared to negative control (heat inactivated cell), medium chain alkanes (C11-C17) showed some reduction in alive *G. kaustophilus* HTA426 (Tamada, 2018).

Careful analysis of the HTA426 genome revealed a candidate gene that could be responsible for the alkane degradation activity. *G. kausthophilus* HTA426 contains a gene identified as ribonucleotide reductase small subunit gene (RNR2, GK2771) in addition to a completed set of RNR gene cluster (RNR1/RNR2, GK0911/GK0912) (~1.8 Mbp apart).

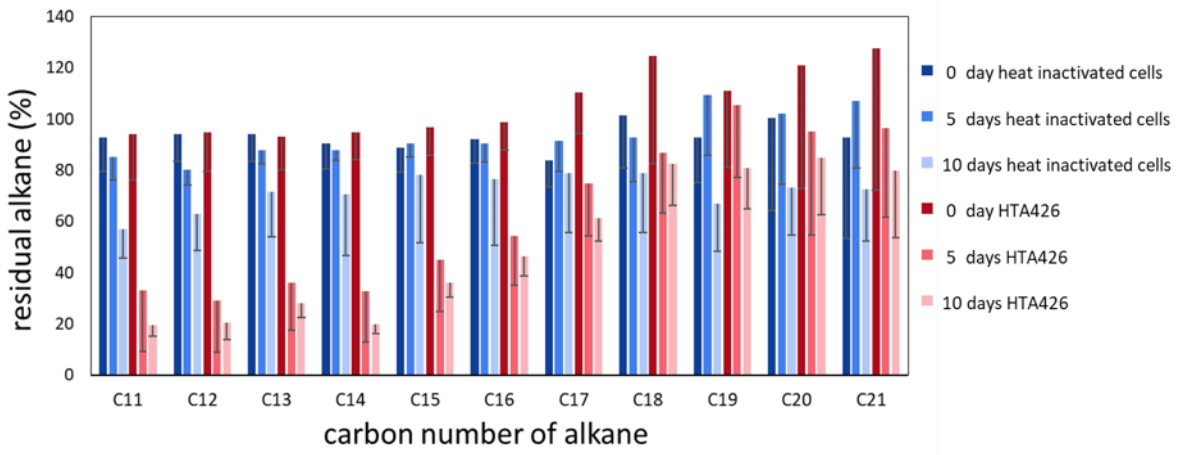


Fig. 4. Alkane degradation by *G. kaustophilus* HTA 426 and heat inactivated (autoclaved) cell at 60℃. The degradation of middle chain alkane can be seen clearly when compared to the inactivated cells. (Tamada, 2018)

0.4 The isolated RNR2, GK2771

As mentioned above we focused our attention on the isolated RNR2 (GK2771). It was hypothesized that GK2771 possibly functions as an alkane degrading enzyme based on the following ideas. RNR2, ribonucleotide reductase small subunit, forms an enzyme with RNR1 large subunit responsible for the conversion of ribonucleotides into their deoxy form. RNR2 has resemblances to one of alkane hydroxylase in term of structure of the active site. A sMMO or soluble methane monooxygenase, capable of oxidizing methane and other short chain gaseous alkane. Both enzymes share a similar diiron/bi metal active site, nearly identical ligand set and together with common working pathway (Fig. 5) (Torrent et al., 2002, Que & Dong,

1996). The isolated RNR2, GK2771, is found to be followed by aldehyde dehydrogenase gene, alcohol dehydrogenase gene and beta oxidation pathway related genes (Fig. 6). The location of GK2771 implies its involvement in alkane metabolism rather than the deoxyribonucleotide synthesis. However, from preliminary experiment, *G. kaustophilus* HTA426 appeared to degrade middle chain alkane which is dissimilar to sMMO substrate.

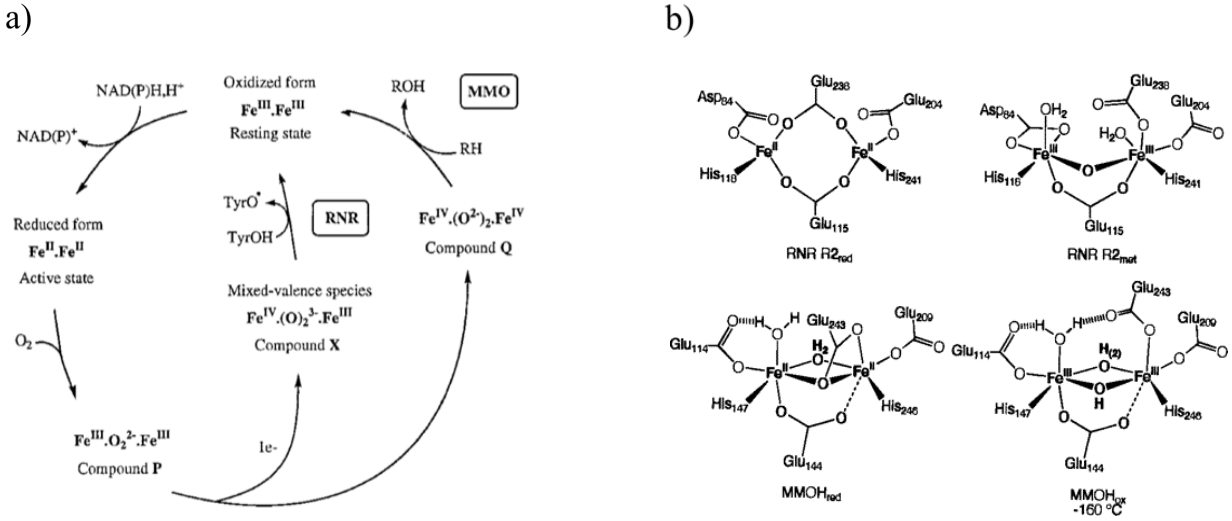


Fig. 5. a) Proposed catalytic cycles for MMO and RNR showing possible common intermediates (Torrent et al., 2002). b) Diiron sites of ribonucleotide reductase, and the hydroxylase component of methane monooxygenase both in reduced (left) and oxidized (right) forms (Que & Dong, 1996). TyrO•:Tyrosine with oxygen radical, TyrOH: Tyrosine with hydroxyl radical. Reprinted with permission from Que & Dong, 1996. Copyright 2024 American Chemical Society.

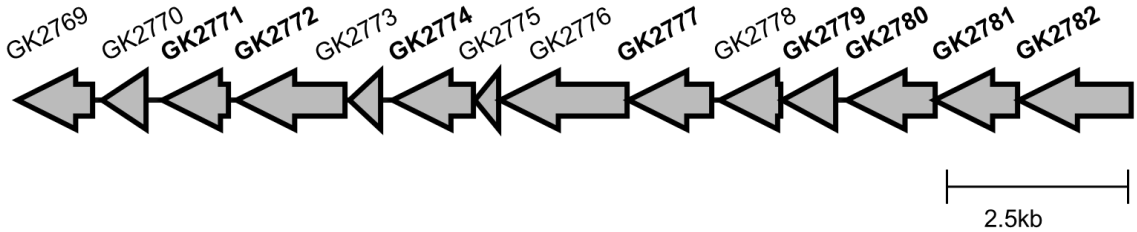


Fig. 6. A gene cluster GK2769 – GK2782 in the *G. kaustophilus* HTA426 genome. GK2769: hypothetical conserved protein, GK2770: hypothetical conserved protein, GK2771: ribonucleotide reductase small subunit, GK2772: NAD-dependent aldehyde dehydrogenase, GK2773: hypothetical protein, GK2774: alcohol dehydrogenase, GK2775: hypothetical protein, GK2776: ATP-binding cassette (ABC) transporter, GK2777: acetyl-CoA acetyltransferase, GK2778: 3-ketoacyl-[acyl carrier protein] reductase, GK2779: enoyl-CoA hydratase, GK2780: acyl-CoA dehydrogenase, GK2781: acyl-CoA dehydrogenase, GK2782: long-chain fatty-acid-CoA ligase

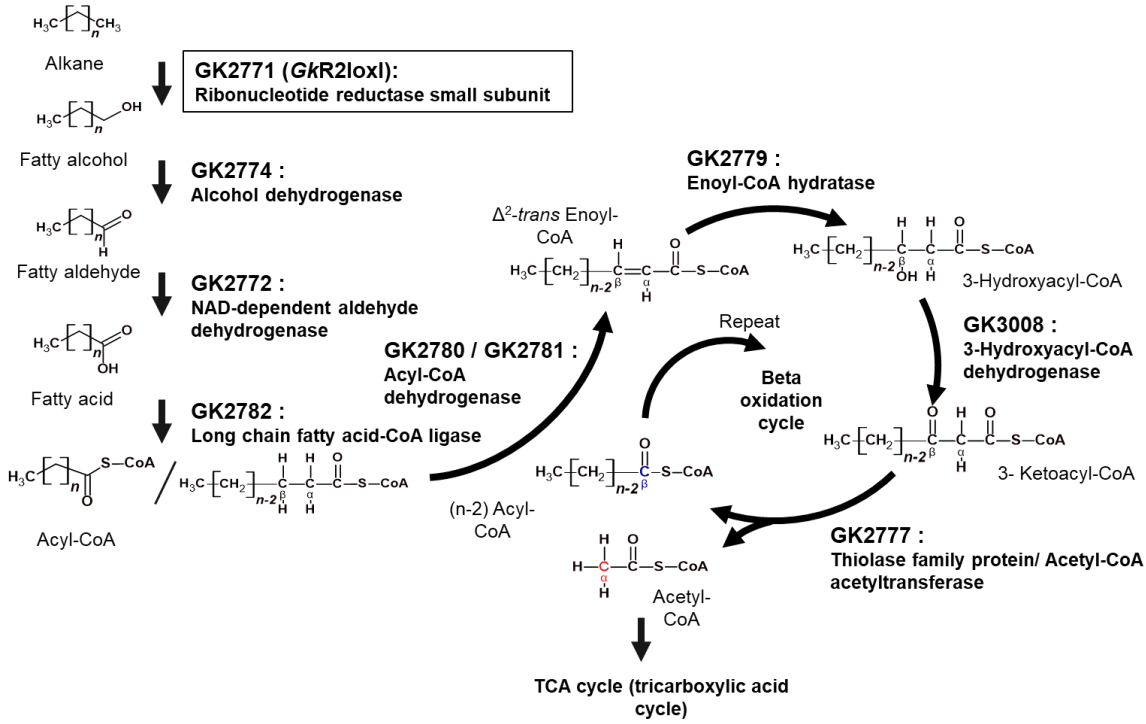


Fig. 7. Estimated alkane degradation pathway of *G. kaustophilus* HTA426 incorporated with gene located nearby the isolated RNR2 (GK2771)

0.5 R2-like ligand-binding oxidases (R2lox)

Initially, a RNR2 from *Chlamydia trachomatis* was identified to lack in forming tyrosine radical, which is an essential for RNR2 activity. The tyrosine was replaced by phenylalanine instead however the protein exhibited an activity to reduce cytidine diphosphate (CDP) to its deoxy form (dCDP) (Roshick et al., 2000), and was proposed to be a new group of Class I RNR, subclass Ic (Hogbom et al., 2004). Later, it had been found that the Ic RNR2 has Mn-Fe active site instead of usual diiron (Voevodskaya et al., 2007; Jiang et al., 2007). Similar Mn-Fe hetero dinuclear protein had been discovered and this led to another new discovery, Anderson and Hogbom 2009, reported a RNR2 Ic like protein from *Mycobacterium tuberculocis* called Rv0233 but had no RNR activity. *M. tuberculocis* contains two more gene homologues of RNR2 except for Rv0233 which belongs to the subclass Ib with dimanganese active site. From phylogenetic perspective, sequence analysis showed that RNR2 Ic like proteins actually form two groups, first group including *C. trachomatis* RNR2 which has RNR activity and another group including the Rv0233. The Rv0233 group does not have conserved C-terminal tyrosine which interacts with RNR1 subunit and essential in *C. trachomatis* RNR2 case. Moreover, some organisms that contain similar protein to Rv0233 lack RNR1 of Class I,

this suggested that the Rv0233 group probably ia not a RNR2 despite similarity in overall sequence and structure. From crystal structure of Rv0233, the metal center is also a Mn-Fe similar to RNR2c, a notable difference from RNR2 protein is that the Rv0233 structure exhibits a large cavity lead directly to the metal site possibly due to changes of conserved radical. It appeared that the space allows substrate or ligand to bind to the metal center like the one observed in BMMs (Bacterial multicomponent monooxygenases) proteins. This led to the author's proposal that these RNR2 subclass Ic-like proteins without the RNR activity should be named as "R2-like ligand binding oxidases". Later GK2771 was identified to be one of the R2lox proteins named Gk2RloxI, the protein crystal structure had been determined (Fig. 8 & Fig. 9) and studied to understand how enzymes with heterodinuclear metal (in this case Mn/ Fe) cofactor distinguish and specify different metal ions in protein assembly. Furthermore, the Mn/Fe cofactor was shown to be capable of catalyzing two electron oxidation reaction similar to Fe/Fe protein of BMMs (Griese et al., 2013). This implies that hydrocarbon oxidation could be one of R2lox protein functions. When purified from *E. coli*, recombinant GkR2loxI protein contains a medium-chain fatty acid (mainly C16 and C18) bound in a hydrophobic tunnel extending from the active site toward the protein surface, MtR2lox (*Mycobacterium tuberculocis* R2lox) also show the same phenomenon (but not all R2lox will be bound by fatty acid) (Griese et al., 2013; Anderson and Hogbom 2009). However, no oxidative and reductive reaction occurred to the bound fatty acid suggested that it is not a substrate but due to a resemblance to substrate in some way hence bounding to the protein. Similar thing happened to ferritin-like enzyme ADO (aldehyde deformylating oxygenases) which was also copurified with a mixture of C16 and C18 fatty acids that are not substrates of the enzyme. ADO catalyzes the conversion of Cn fatty aldehydes into the corresponding Cn–1 alkanes or alkenes (Mak et al., 2020). Even so, aldehyde is also not a substrate for R2lox. Currently, the function and substrate of R2lox protein are still unknown (Diamanti et al., 2022).

In this dissertation, I demonstrate that *G. kaustophilus* HTA426 could indeed degrade medium to long chain alkane, despite the absence of a conventional alkane hydroxylase gene in its genome. Moreover, the hypothesis that GK2117 is a novel type of alkane monooxygenase was proved by showing recombinant GK2771 enzyme prepared in *E. coli* BL21 (DE3) exhibits the activity of alkane degradation. In addition, an activity assay was conducted on GK2772, an aldehyde dehydrogenase gene, to further support the claim of GK2771's function as an alkane degradation gene.

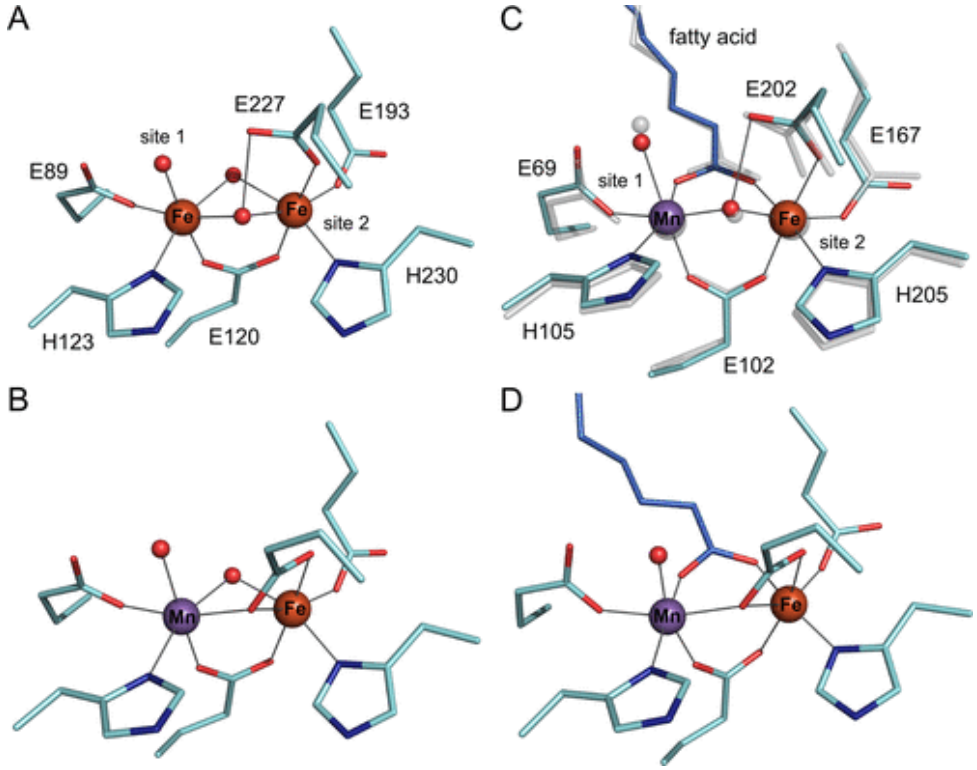


Fig. 8. Crystal structures of R2c and R2lox protein active site demonstrating their similarity. (A) The oxidized Fe/Fe-cofactor of CtR2c (*Chlamydia trachomatis* ribonucleotide reductase small subunit) at 1.7 Å resolution. (B) The reduced Mn/Fe cofactor of CtR2c at 1.8 Å resolution. (C) Superposition of the oxidized Mn/Fe-bound active sites of GkR2loxI (cyan, metal-coordinating residues; blue, fatty acid ligand) and MtR2lox (R2lox of *Mycobacterium tuberculosis*, Rv0233), both at 1.9 Å resolution. Residue numbering is for GkR2loxI. (D) The reduced Mn/Fe-bound active site of GkR2loxI at 1.9 Å resolution. (Shafaat et al., 2014)

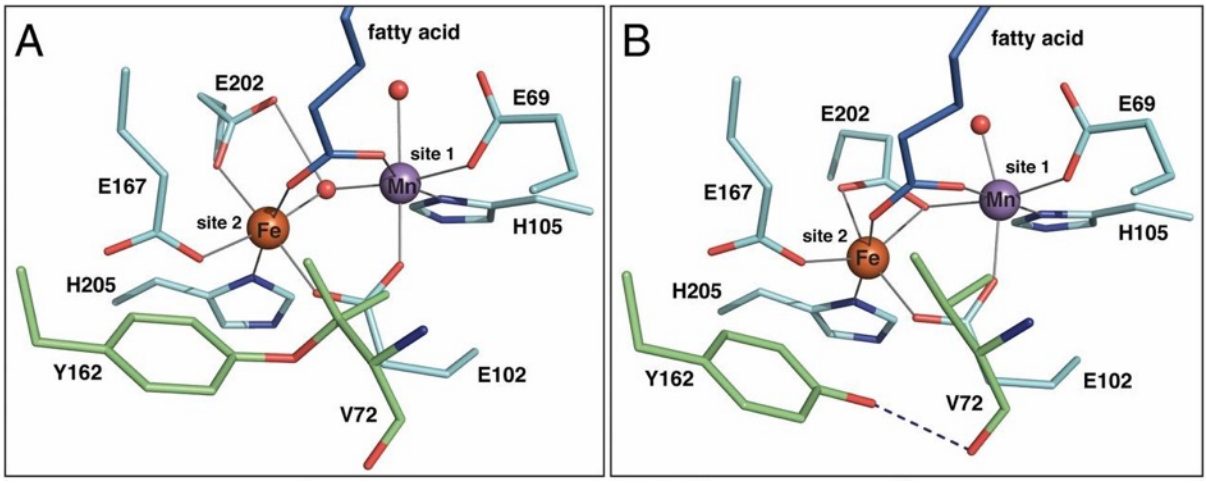


Fig. 9. Crystal structures of *Gk*R2loxI active site showing unique tyrosine-valine crosslink. (*A*) The oxidized Mn/Fe-bound active site at 1.9 Å resolution. (*B*) The reduced Mn/Fe-bound active site at 1.9 Å resolution (Griese et al., 2013)

CHAPTER I

Alkane degradation by *Geobacillus kaustophilus* HTA426

1.1 Materials and methods

1.1.1 Growth of *G. kaustophilus* HTA426 on minimal medium

G. kaustophilus HTA426 colony from LB agar plate was transferred to 5 mL LB and cultured at 60℃, 120 rpm shaking overnight as a pre-culture. The LBM minimal media (Kato et al., 2001) main culture was inoculated by 1% v/v from the pre-culture. Then 0.1% v/v of standard oil (Tokyo chemical), a mixture of alkanes, was added as the sole carbon source, glucose as a carbon source and the media without pre-culture inoculation were used as positive and negative control, respectively. The main cultures were incubated similarly to the pre-culture's condition. OD_{600} was observed at 0 hr and every 24 hr afterwards.

1.1.2 Alkane degradation test

The bacteria were pre-cultured in 20 mL of LB at 60℃, 120 rpm overnight followed by 1% inoculation into appropriate amount LB and further incubated overnight. Cells were harvested and washed twice with fresh LBM, then resuspended in LBM to get final OD_{600} around 5.00. The culture was transferred into sterile 20 mL glass vial (Maruemu Corporation), 1.5 mL each with 1.5 μL (0.1% v/v) standard oil and closed with butyl rubber and aluminum cap. All samples were incubated at 60℃ in a standing condition for 0, 1, 3, 7, 9 and 11 days. The whole culture in the vial was extracted after each incubation day by 1.5 mL of acetone:hexane (1:1), at an equal volume to the sample, which contained 40 ppm biphenyl as an internal standard for Gas chromatography analysis. Solvent extraction was performed by shaking with Strong shaker SR-2 DS (TAITEC) at 300 rpm for 20 min. The extracted samples were left stand at room temperature for 24 hr, then the upper solvent layer was used to measure a remaining amount of alkane by GC/FID (GC-2014, Shimadzu). The relative amount of each alkane was calculated by comparing the peak area with biphenyl. All experiments were prepared in triplicate glass vials. Blank or negative control is LBM without bacteria.

The relative amount of alkane in the sample was calculated by comparing a peak area with biphenyl.

$$\text{Remaining alkane (\%)} = \frac{\text{relative amount of alkane in the sample}}{\text{relative amount of alkane at 0 hr of pararellely prepared sample}}$$

1.1.3 Gas chromatography

For GC/FID (GC-2014, Shimadzu) analysis, a volume of 1 µl from the upper layer of the extracted sample was used. The GC system was equipped with a flame ionization detector and a HP-5 nonpolar capillary column (30 m length, 0.32 mm diameter, 0.25 µm film thickness). Helium was used as carrier gas at 25 ml/min flow rate. The temperature program involved increasing the column oven temperature from 50˚C to 300˚C and holding it at 300˚C for 10 min.

1.2 Results

1.2.1 Growth of *G. kaustophilus* HTA426 on minimal medium

After 1 day incubation of *G. kaustophilus* HTA426 in LBM + 0.1% glucose culture became turbid but no change was observed for negative control and *G. kaustophilus* HTA426 in LBM standard oil supplemented sample. The *G. kaustophilus* HTA426 in LBM supplemented with standard oil remained clear and no cell growth was observed even after 5 days. This suggests that alkane does not support the cell growth of HTA426 as a sole carbon source.

1.2.2 Alkane degradation test

The alkane degradation test on mineral medium LBM was done twice. In the first experiment, incubation period was for 5 days (Fig. 10). After 5 days of incubation, the total amount of all alkanes (C10 – C24) in *G. kaustophilus* HTA426 samples was reduced to roughly 10%. However, due to high temperature and long incubation time, a reduction of alkane was also observed in blank sample, the amount of alkanes reduced to around 20% - 40% except for C10 (57%), C23 (96%) and C24 (100%).

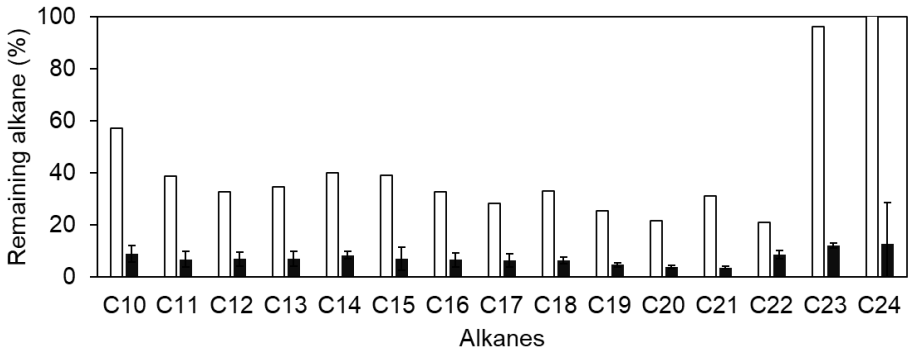
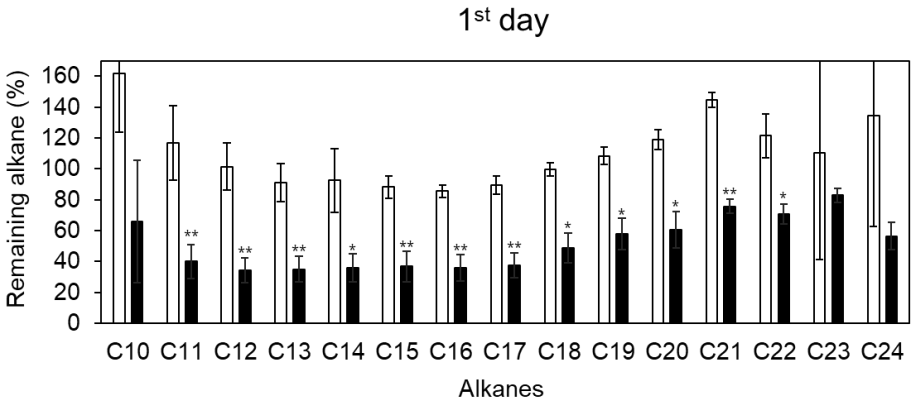
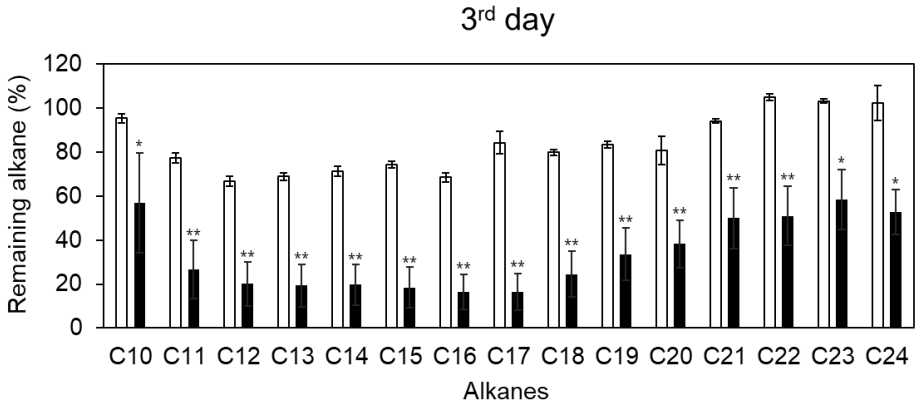
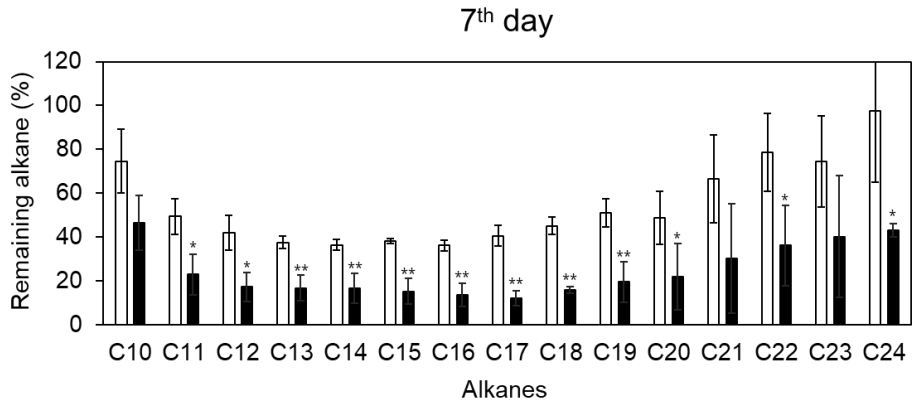


Fig. 10. Alkane degradation of *G. kaustophilus* HTA426 in LBM containing 0.1% (v/v) standard oil, 5 days at 60℃. White bar, Blank (media only); black bar, *G. kaustophilus* HTA426

In the second experiment, *G. kaustophilus* HTA426 was incubated and observed at two days interval for 11-days to follow alkane degradation pattern (Fig. 11). Samples on 1st, 3rd, 7th, 9th, and 11th day were collected and analyzed by GC/FID. Degradation of most alkanes started on the first day of incubation, while alkanes in blank still remained at 100%. *G. kaustophilus* HTA426 reduced middle chain alkane (C11-C18) to 40% and shorter and longer alkanes were reduced to around 60% - 80%. The reduction of alkane continued over the course of incubation time for both HTA426 and blank samples. The reduction in the blank sample started from the 3rd day, C11 – C20 around 70 – 80% while other alkanes were still at 100% remaining. At the end of experiment (11th day) C10 – C17 were around 30 – 40% and C18 – C24 were 50 – 90% remained. For *G. kaustophilus* HTA426, the alkane degradation seemed to stop after the 7th day due to the sample having the lowest amount of alkane left, C11 – C20 were reduced to 10 – 20% of the initial amount, C16 and C17 are the lowest at 13.5% and 12.0% respectively. Remarkably, the remaining percentage of C18 – C23 on 9th day sample were lower than C11 – C17 which differed from earlier samples. This may suggest that the bacteria at later day switched to utilize more C18 – C23 after used up C11 – C17 for certain amount.


1st day
Remaining alkane (%)
0
20
40
60
80
100
120
140
160
C10 C11 C12 C13 C14 C15 C16 C17 C18 C19 C20 C21 C22 C23 C24
Alkanes


3rd day
Remaining alkane (%)
0
20
40
60
80
100
120
C10 C11 C12 C13 C14 C15 C16 C17 C18 C19 C20 C21 C22 C23 C24
Alkanes


7th day
Remaining alkane (%)
0
20
40
60
80
100
120
C10 C11 C12 C13 C14 C15 C16 C17 C18 C19 C20 C21 C22 C23 C24
Alkanes

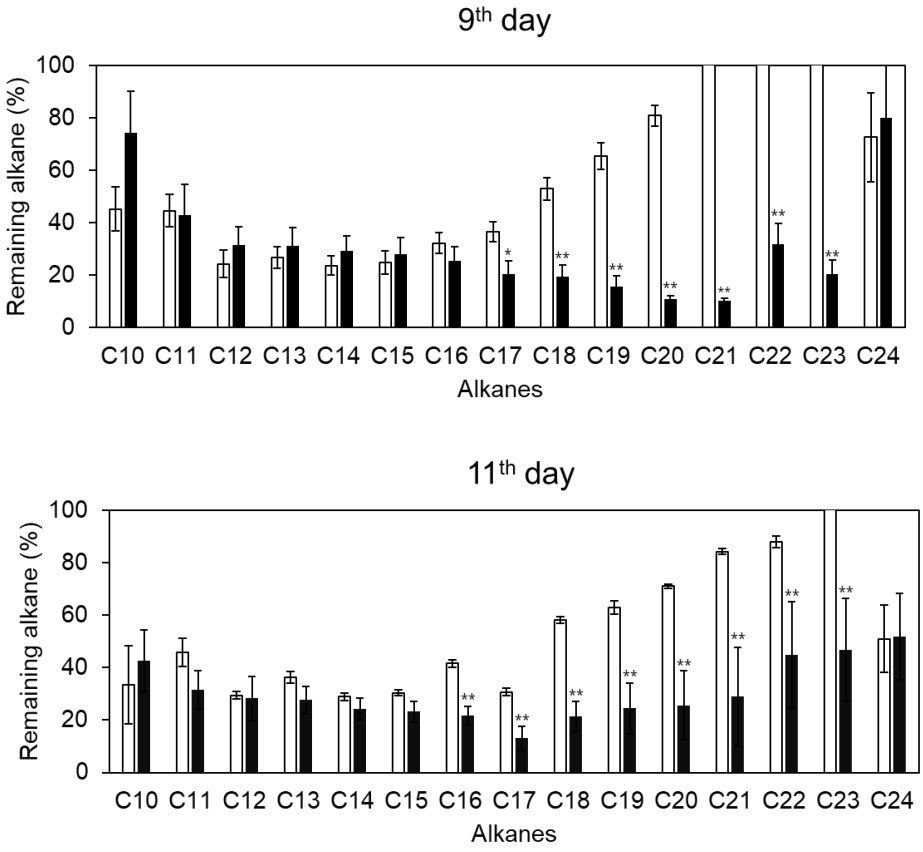


Fig. 11. Comparison of the remaining amounts of alkanes between vials containing *G. kaustophilus* HTA426 cells and the medium only control over 11 days period. Alkane degradation of HTA426 in LBM containing 0.1% (v/v) standard oil, at 60℃. White bar, Blank (media only); black bar, *G. kaustophilus* HTA426, Single asterisk: $P < 0.05$, Double asterisk: $P < 0.01$

1.3 Discussion

The results clearly demonstrate that *G. kaustophilus* HTA426 is able to degrade C10 – C24 alkane. Although the bacteria could not grow on the alkanes as sole carbon source in synthtetic LBM medium. It cannot deny an idea that some essential trace elements for the growth of HTA426 are missing in LBM. Compared to previous data on alkane degradation of *G. kaustophilus* HTA426 in LB (nutrient rich condition) the degradation pattern was different. In LB, over 5 days of incubation the *G. kaustophilus* HTA426 degraded only C11 – C17 but in the minimal medium condition of LBM, the bacteria can degrade wider range of alkane (C10 – C24) (Fig. 12). This finding supports the phenomenon that *G. kaustophilus* MK93 degraded

C20 was valid. In addition, it is suggested that alkane degrading enzymes of HTA426 such as alkane monooxygenase, alcohol dehydrogenase, and aldehyde dehydrogenase prefer medium hain substrates rather than shorter or longer aliphatic substrates.

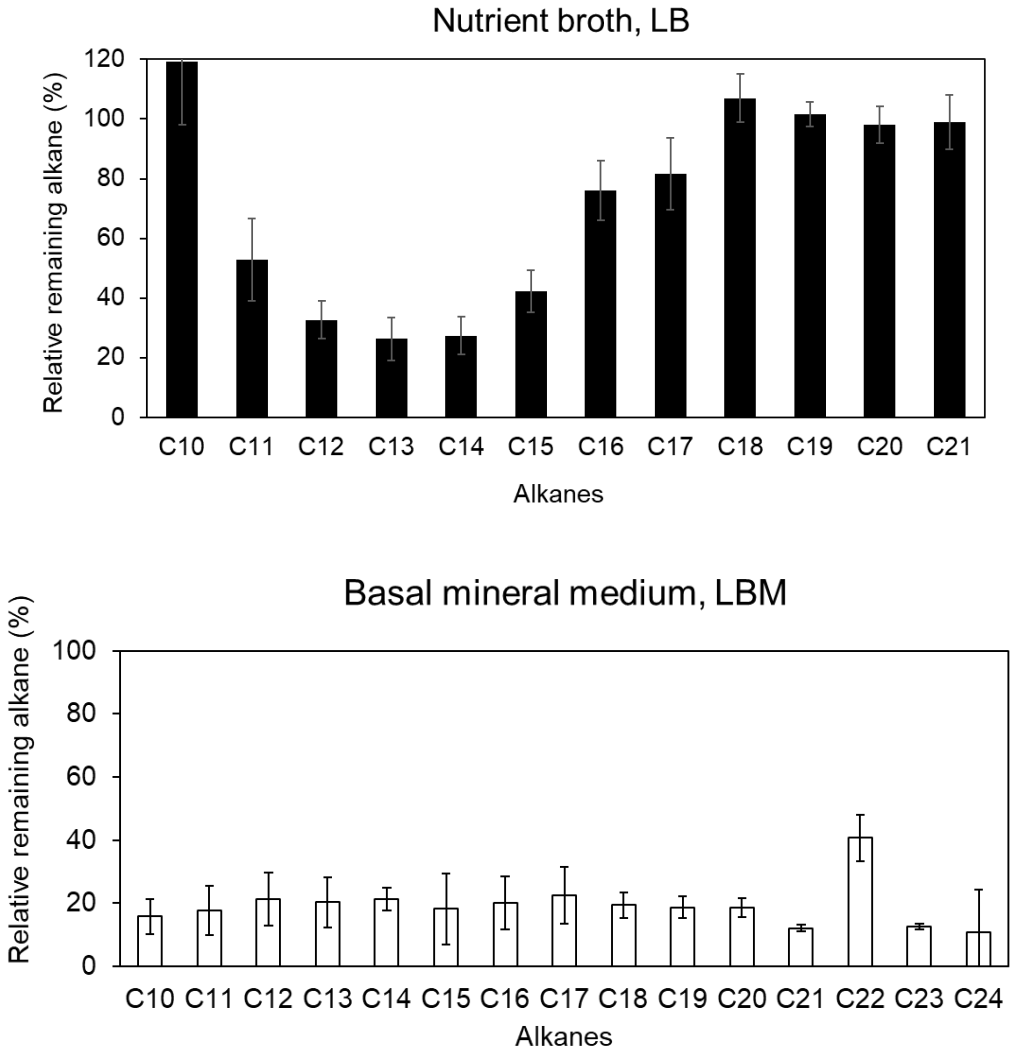


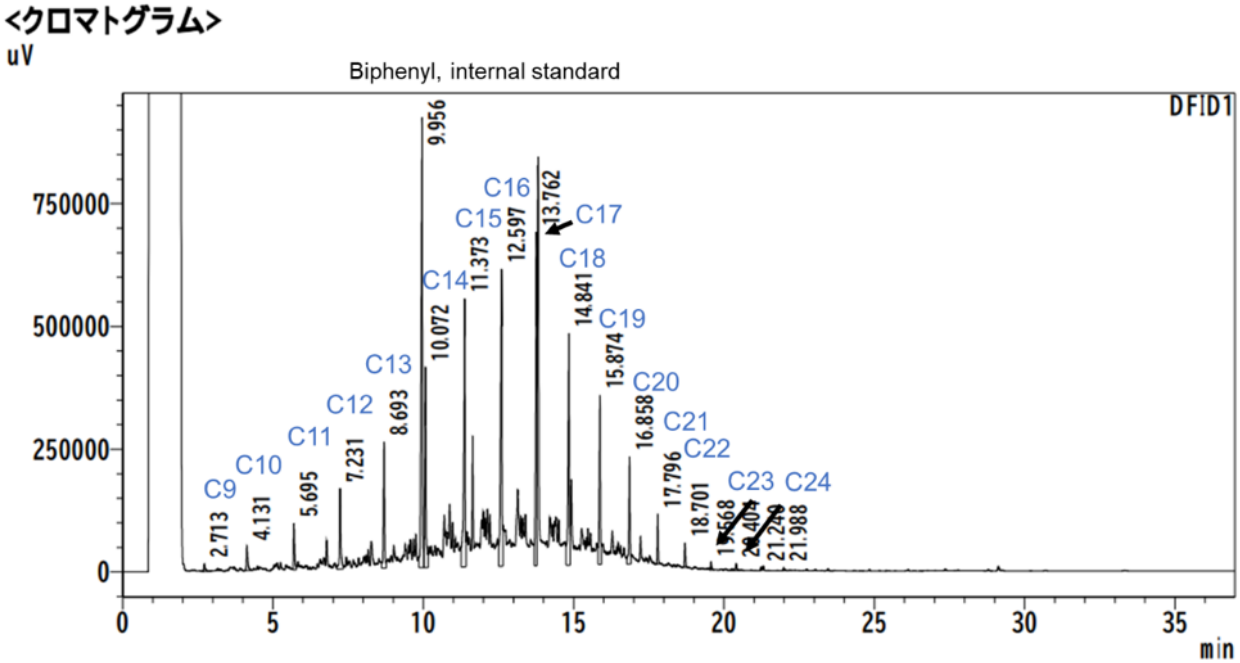
Fig. 12. Comparison of the relative remaining amounts of alkanes between *G. kaustophilus* HTA426 culture in LB and LBM media after 5 days at 60˚C. Each medium contained 0.1% (v/v) standard oil. Relative remaining alkane (%) was the amount of remaining alkane in *G. kaustophilus* HTA426 divided by the amount of alkane remained in blank of the same day.

The result of the 11 days incubation experiment provides an insight into the alkane degradation pattern of *G. kaustophilus* HTA426. Moreover, the 1st day result clearly indicates that the reduction of alkane was due to *G. kaustophilus* HTA426 not from non-biological factor influence including evaporation and adsorption to the rubber cap. No reduction of alkane by autoclaved HTA426 cells were also confirmed (data not shown). The results of 1st, 3rd, 7th (5th day result could not be obtained due to technical error) have a similar degradation trend, all of

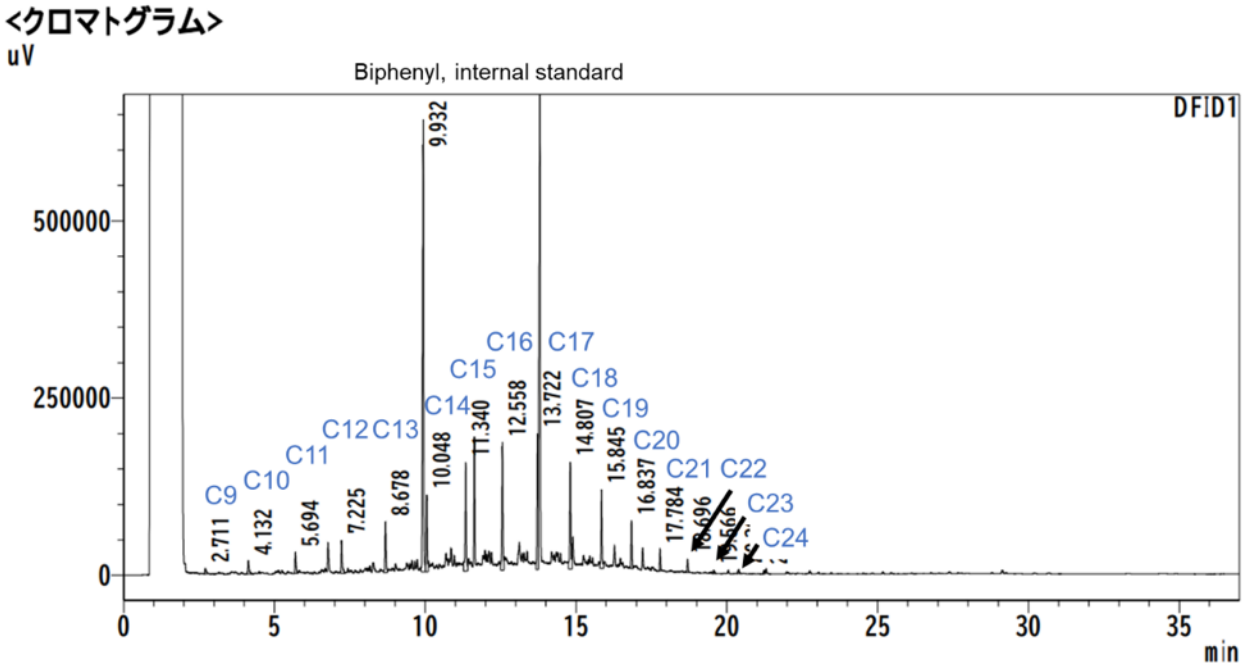
alkanes were reduced but the most degraded portion was medium chain alkane (C11 – C17). While the degradation from *G. kaustophilus* HTA426 appeared to stop since the day 7th due to no further reduction was observed and the decreasing of C10 – C17 alkane in blank samples caught up to the bacteria sample. As mentioned in the 1.2, the C18 – C23 of the *G. kaustophilus* HTA426 9th day sample were lower than usual. Also, the decrease in blank sample on the longer chain alkane was lesser than the shorter chain alkane, resulting in this occurred from the bacteria. It was observed again that *G. kaustophilus* HTA426 prefers medium chain alkane (C11 – C18) over longer chain alkane (> C18) and shorter chain alkane (C10) (3rd day, Fig. 11).

It is worth noting that the quantity of each alkane present in the standard oil mixture is unequal as shown in Fig. 13 a), the quantities of each alkane type in the standard oil sample exhibit a distribution pattern similar to a normal distribution curve. C15 – C18 are the most prevalent alkane type (~50% of the total alkane). It is also possible that *G. kaustophilus* HTA426 dominantly degraded C15 – C18 due to higher amount in the substrate compared to the other alkane.

a)



b)



c)

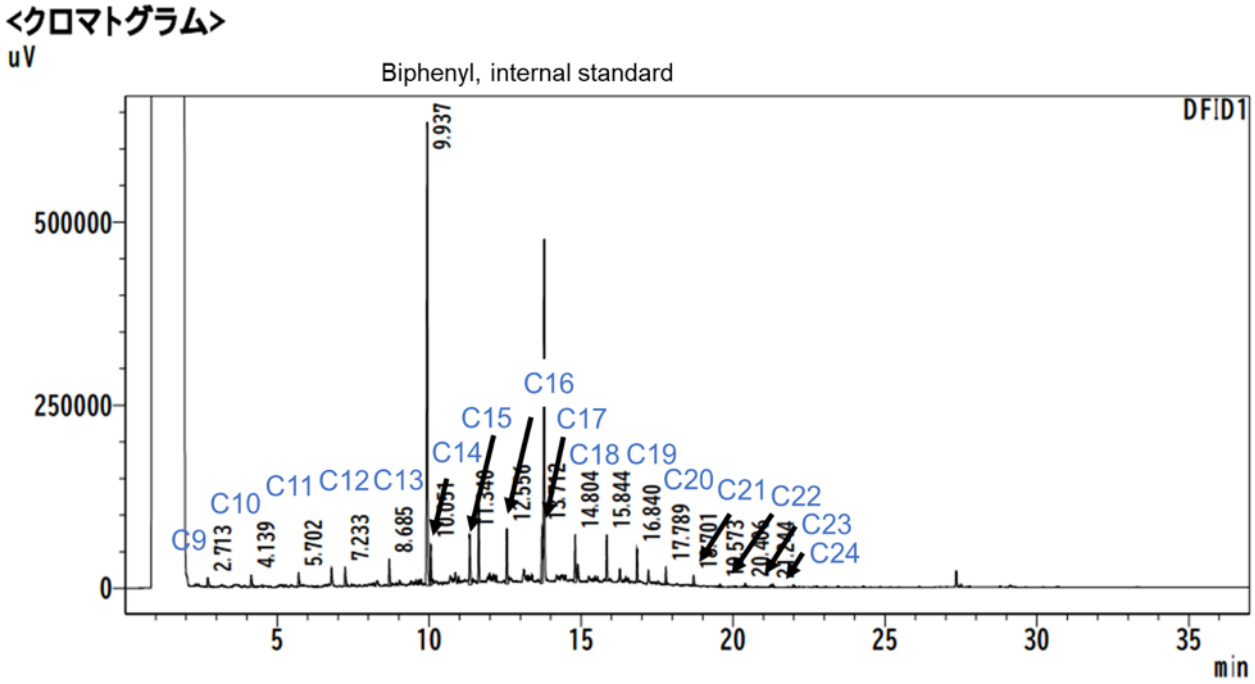


Fig. 13. Chromatogram from GC/FID analysis of alkane degradation test samples at 60℃. a) *G. kaustophilus* HTA426 cell in LBM, 0 day; b) blank, LBM 11 days; c) *G. kaustophilus* HTA426 cell in LBM, 11 days. Peaks at 9.9 min and 13.7 min are of Biphenyl, internal standard and unknown impurity compound.

CHAPTER II

Heterologous gene expression of GK2771 gene in *E. coli*

2.1 Materials and methods

2.1.1 Strains and plasmids

The heterologous expression of the GK2771 gene, has been done in *E. coli* BL21 (DE3) using pET28a as an expression vector containing strong T7 promoter. *E. coli* BL21 (DE3) is a strain constructed for a high and strictly controlled recombinant protein expression system using T7 RNA polymerase gene expression system with compatible pET expression vectors. The bacterial strains and plasmids used or constructed in this study are listed in **Table 2.**

Table 2. Strains and plasmids used in heterologous gene expression of GK2771 gene (RNR2) in *E. coli* BL21 (DE3).

Strain or plasmid		Reference
Strains		
E. coli DH5a	Host strain for pUC19 cloning vector F$^-$ *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3)	-
E. coli BL21 (DE3)	Host strain for pET gene expression vector Phage lamdaDE3 lysogenic strain containing T7 RNA polymerase gene under the control of lac promoter operator	pET system manual, Novagen
Cloning and expression vectors		
pUC19	Cloning vector, containing Ap^r gene	Yanisch-Perron *et al.,* 1985
pET28a	T7 promoter/ lac operator, containing Kan^r gene and lacI	pET system manual, Novagen
pUC19-GK2771	pUC19 with GK2771 gene	This study
pET28a-GK2771	pET28a with GK2771 gene	This study
	pUC19 with GK2771 gene + His tag	This study
pET28a-GK2771 His	pET28a with GK2771 gene + His tag	This study

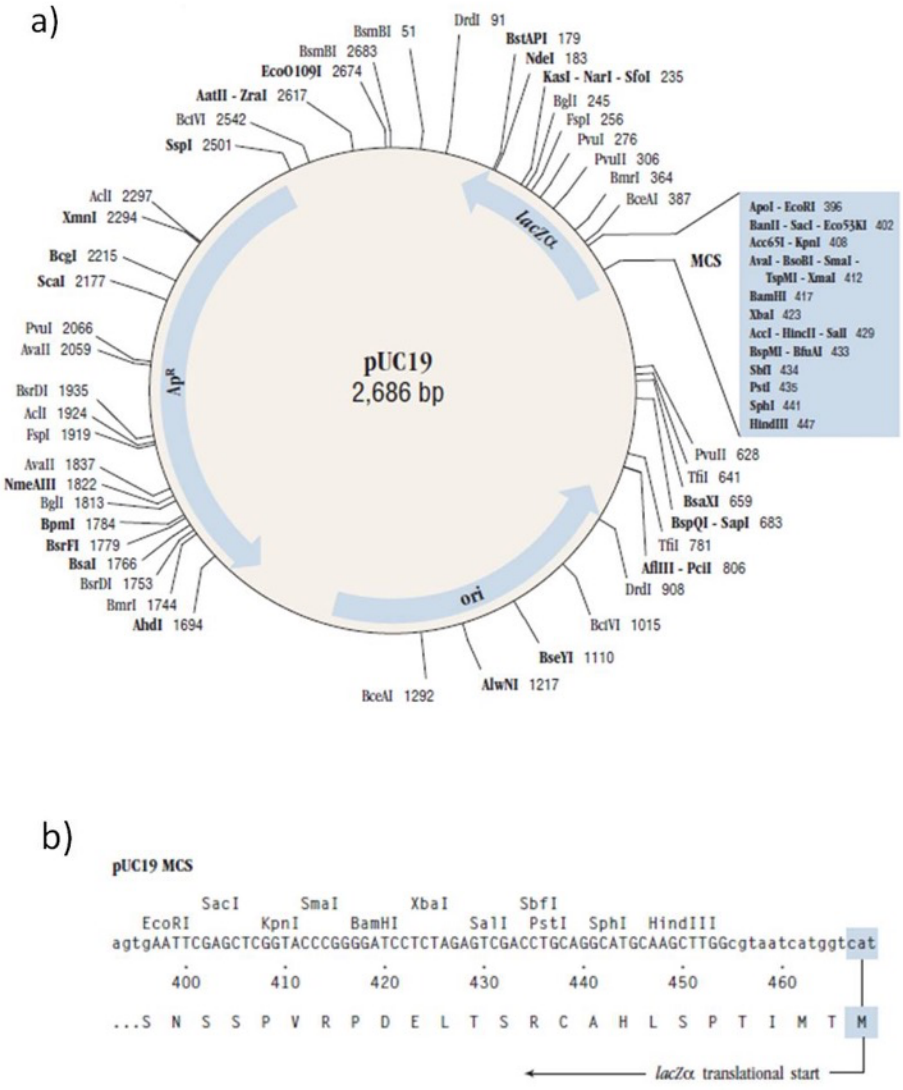


Fig. 14. a) Physical map of pUC19 and b) the multiple cloning sites (New England BioLabs inc.)

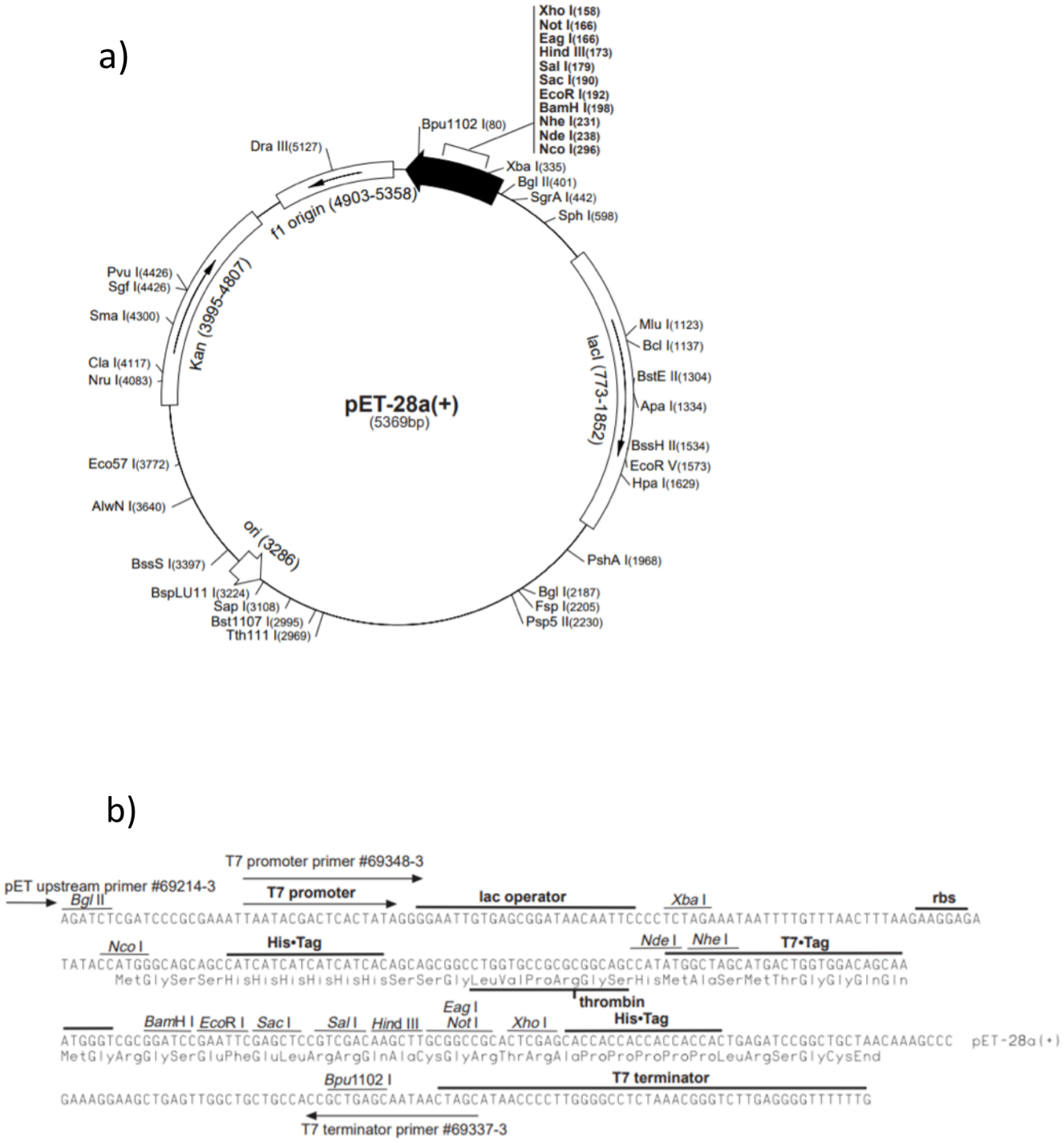


Fig. 15. a) Physical map of pET28a and the b) multiple cloning sites (Novagen)

2.1.2 Cloning of GK2771 gene and construction of gene expression plasmid

2.1.2.1 *gk2771*

Genomic DNA of *G. kaustophilus* HTA426 was prepared by InstaGene^{TM} Matrix (Bio-Rad). The GK2771 gene was amplified by PCR with KOD-Plus-Neo DNA polymerase (Toyobo, Kyoto, Japan). The primers, F71 (*Nco*I) and R71 (*Eco*RI) used for PCR amplification were designed by attaching *Nco*I and *Eco*RI restriction sites to the forward start and the reverse end of GK2771 gene sequence, respectively. The PCR amplified GK2771 gene with blunt end was cloned into *Sma*I site of pUC19. The heat shock method (Inoue et al., 1990) was used for *E. coli* DH5α transformation. The insert direction and no sequence error were confirmed by restriction enzyme digestion and sequencing, respectively. Next, the GK2771 gene fragment was intended to clone into *Nco*I and *Eco*RI restriction sites of pET28a vector for expression but due to unavailability to obtain desired transformant for long period of time, the strategy was changed to blunt end ligation instead. The insert from pUC19-GK2771 could not be used because of additional nucleotide at filling up the *Nco*I site would hamper the gene translation (Fig.14 c.) Thus, new forward primers, F71+atg and F71-atg, were used to directly amplified GK2771 gene and cloned into filling up *Nco*I site of pET28a vector except additional start codon (ATG) in case of F71+atg primer. The new PCR products with blunt end were purified by QIAquick PCR Purification Kit (Qiagen) for ligation. pET28a digested with *Nco*I enzyme (sticky end) was filled up by DNA blunting kit (Takara). After the new PCR amplified GK2771 gene fragment was lagated with blunted *Nco*I site of pET28a vector, the product pETa-GK2771 was transferred into *E. coli* DH5α. The sequence with no error was confirmed and subsequently transferred into *E. coli* BL21 (DE3).

Table 3. Primers used in heterologous gene expression of GK2771 in *E. coli* BL21 (DE3)

Genes	Length (bp)	Primer name	Sequences	Restriction site
gk2771	909	F71 (*Nco*I)	5′ CATGCCATGGATGGTCCATCATGACGG 3′	5' *Nco*I
		R71 (*Eco*RI)	5′ CGGAATTCTCATGACTCAGCAGCC 3′	-*Eco*RI 3'
		F71+atg	5′ **ATG**GTCCATCATGACGG 3′	
		F71-atg	5′ GTCCATCATGACGG 3′	

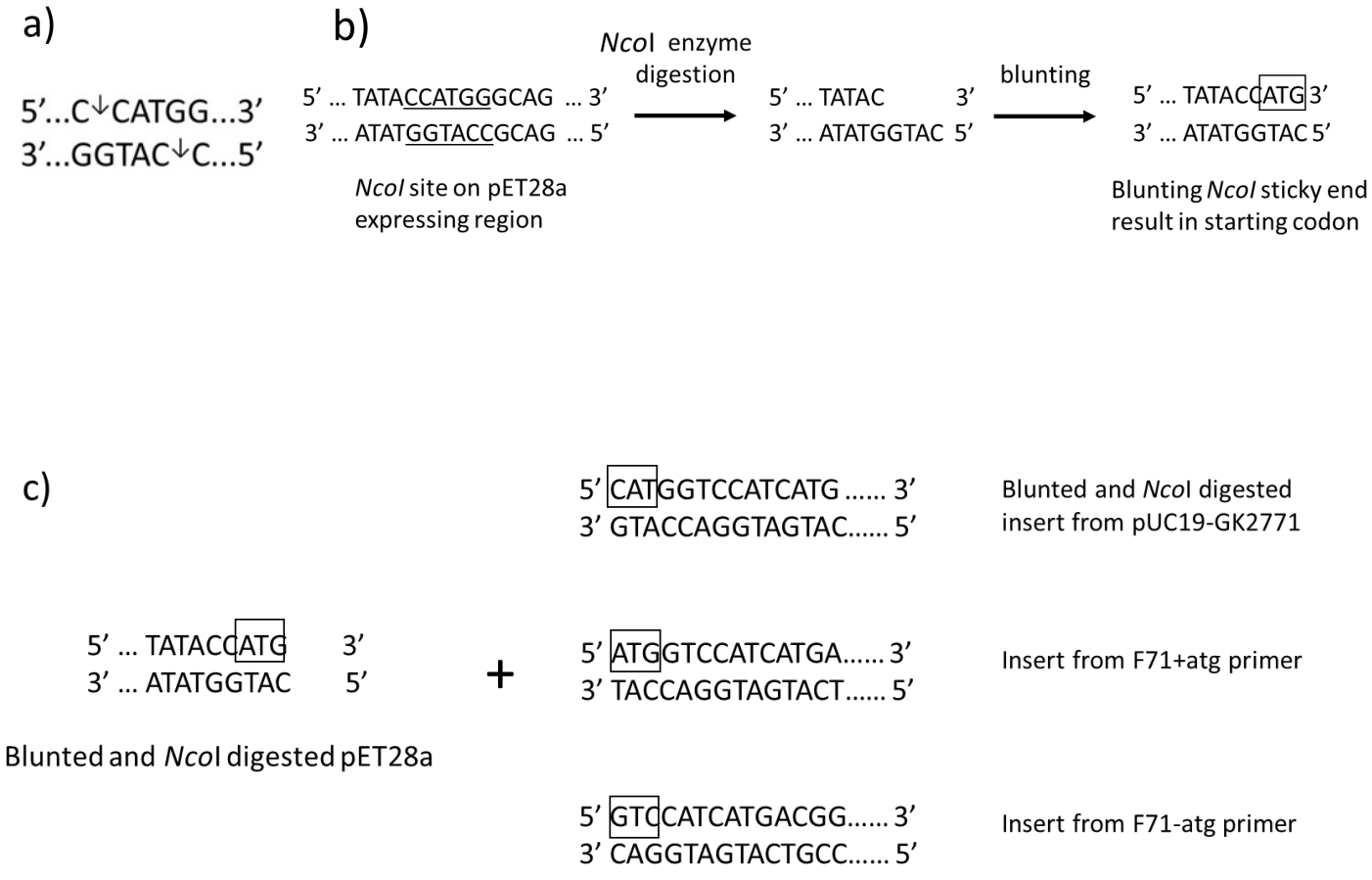


Fig. 16. a) Cutting site of *Nco*I restriction enzyme, b) DNA Blunting of *Nco*I digested pET28a, the reason why a new forward primer was required to amplify the insert for blunt end ligation, c) Blunt end ligation of blunted *Nco*I digested pET28a and different inserts. Top) Blunt ended insert DNA from pUC19-GK2771 brings additional cytosine which causes a frameshift.

2.1.2.2 Mutated *gk2771*

The mutant was designed to lower the protein alkane degradation activity by altering the amino acids that connect to an iron atom at center of the protein. The glutamic acid (202) and Histidine (205) were converted to alanine by KOD – Plus mutagenesis kit (Toyobo) which utilizes inverse PCR technique. Following the kit instruction, the pET28a – GK2771 as a template and, inact 2771F and inact 2771R primers were used for site direct mutation inverse PCR. *Dpn*I was added into the PCR product to digest template plasmid, then ligation mixture was added to perform a self-ligation of the amplified mutated pET28a – GK2771 plasmid fragment. The ligated plasmid was transferred into *E. coli* DH5α. The sequence with no error was confirmed and subsequently transferred into *E. coli* BL21 (DE3).

Table 4. Primers used in mutated pET28a - GK2771 construction

Genes	Length (bp)	Primer name	Sequences	Restriction site
gk2771	909	F Inact 71	5′ CATGGACGCAGGGCGGGCCATTCAGTTCGG 3′	
		R Inact 71	5′ CTGAATGGCCCGCCCTGCGTCCATGTTC 3′	



Fig. 17. The change in amino acid of GK2771 mutant, the glutamic acid (E202) and Histidine (H205) were converted to alanine.

2.1.2.3 *gk2771* with His tag

Due to the difficulty of obtaining purified GK2771, a new forward primer was designed to amplify GK2771 fragment with NdeI restriction site to utilize His tag of pET28a. Similar to the wildtype GK2771 except for the NdeI F71 primer, new GK2771 with NdeI – EcoRI restriction site was amplified and cloned into SmaI site of pUC19 and transferred into *E. coli* DH5α. The presence of insert was confirmed by restriction enzyme digestion and sequencing, respectively. The NdeI – EcoRI digested insert fragment was ligated to NdeI – EcoRI digested pET28a plasmid. The ligation product pET28a - GK2771 His tag was transferred into *E. coli* DH5α. The sequence with no error was confirmed and subsequently transferred into *E. coli* BL21 (DE3).

In addition, the mutated GK2771 His tag was constructed in similar fashion to 2.1.2.2 while using pET28a - GK2771 His tag as a template.

Table 5. Primers used in pET28a - GK2771 His tag construction

Genes	Length (bp)	Primer name	Sequences	Restriction site
gk2771	909	F71His	5′ GACTGCATATGGTCCATCATGACGG 3′	*Nde*I
		R71	5′ CGGAATTCTCATGACTCAGCAGCC 3′	*Eco*RI

2.1.3 Heterologous gene expression in *E. coli*

E. coli BL21 (DE3) harboring plasmid contained pET28a - GK2771 or other variants was pre-culture in small volume of LB at 37℃, 120 rpm overnight. The culture was transferred to main culture by 1% inoculum size and continued incubation under the same condition until reaching mid log phase or an OD_{600} is 0.4 - 0.5 (approximately 2-3 hour). Subsequently, IPTG was added to final concentration 0.1 mM, and the culture was further incubated for an additional 4 hours. The whole culture was collected and centrifuged at 13,000 rpm for 10 min 4˚C, followed by removal of the supernatant. The cell pellet was resuspended in an appropriate buffer for further experiment and small volume aiming for high concentration of the protein. The resuspended pellet was sonicated using the Handy Sonic model CR-20P (TOMY Seiko co., ltd, Tokyo, Japan) on ice to prevent protein denaturation (4 cycles of 30 seconds sonication followed by 30 seconds resting period). The sonicated samples were centrifuged at 13,000 rpm, 10 min at 4˚C to discard the cell pellet. A portion of the sample was analyzed by SDS-PAGE to confirm the protein production. *E. coli* BL21 (DE3) contained the pET28a vector without insert and non-induced culture were used as negative controls.

2.1.4 Protein purification

2.1.4.1 GK2771 and mutated GK2771

Two types of columns chromatography were used for the GK2771 without histidine tag variants, anion exchange chromatography and hydrophobic interaction chromatography. The chromatography purification was performed at 1 ml/min flow rate, at room temperature and protein was detected at 280 nm wavelength, which purification condition are as follow:

1. Anion exchange chromatography
   Column: HiTrap Q FF 1 mL (Cytiva)
   Binding buffer (Buffer A) : 20 mM Tris-HCl, pH 8.0
   Elution buffer (Buffer B) : 20 mM Tris-HCl, 1.0 M NaCl, pH 8.0

2. Hydrophobic interaction chromatography
   Column: HiTrap Octyl FF 1 mL (Cytiva)
   Binding buffer (Buffer A): 50 mM sodium phosphate, 1.5 M ammonium sulphate, pH 7.0
   Elution buffer (Buffer B): 50 mM sodium phosphate, pH 7.0

After removal of sonicated cell, the lysate was heated at 70˚C for 10 min before injection to the column. Subsequently, 1 mL of the lysate was loaded into the column, the purification started with 10 min of buffer A, followed by 20 min of elution gradient from 0% to 100% of buffer B. Buffer B was continued to flow at 100% for 10 min and the column was washed by buffer A for another 10 min. Throughout the purification, eluted protein fractions were collected corresponding to peaks that appeared on the monitor. Due to high salt concentration in the collected sample, acetone precipitation was conducted before SDS page analysis. First, one volume of the eluted sample was mixed with four volumes of -20˚C acetone and vortexed. The mixture was then incubated at -20˚C for a minimum of 1 hour or left overnight. Following incubation, the samples were centrifuged at 13,000 rpm for 10 min, resulting in the appearance of the protein pellet. The supernatant was removed, and the pellet was air-dried at room temperature for 30 min. Finally, the pellet was dissolved in 5μL to 10 μL of sterile water for SDS PAGE analysis.

2.1.4.2.1 His tagged GK2771

For the His tagged recombinant protein, an immobilized metal ion affinity chromatography (IMAC) column for histidine-tagged protein purification was used, which purification condition are as follows:

Column: HisTrap HP 1 ml (Cytiva)

Binding buffer (Buffer A): 20 Mm sodium phosphate, 0.5 M NaCl, 40 Mm imidazole, pH 7.4

Elution buffer (Buffer B) : 20 Mm sodium phosphate, 0.5 M NaCl, 500 Mm imidazole, pH 7.4

The purification process was similar to 2.1.4.1 but the elution gradient was 15 min. After the GK2771 – His tag protein's peak was identified, the elution method was changed from gradient to step wise elution. The gradient was replaced by 10 min each of following buffer B concentration 0%, 40%, 100% and 0% for total 40 min. The eluted protein fractions underwent desalting by PD-10 Desalting column (Cytiva) to replace the elution buffer with 0.1 M Tris buffer, Ph 7.0. The protein concentration was measured by Bio-rad Protein Assay (Bio-rad) and the protein bands were analyzed on 15% SDS-PAGE.

2.1.5 Alkane degradation assay

For recombinant enzyme, after the recombinant protein was obtained in form of crude enzyme or purified enzyme, the concentration was measured by Bio-Rad protein assay (Bio-rad). The crude enzyme was transferred to sterile 20 ml vial (Maruemu coporation), then NADH, $FeSO_4$ and 1.5 µl (0.1% v/v) of standard oil (Tokyo Chemical Industry) were added. The final volume was 1.5 ml per vial and the concentration was 10 mg/ml for crude enzyme and 0.1 Mm for NADH and $FeSO_4$. The vials were sealed by butyl rubber and aluminium cap. All samples were incubated at 37˚C, standing condition. Every condition was done in triplicate.

After reaching the designed incubation time, the whole culture was extracted by adding 1 volume of hexane:acetone (1:1) which contained 40 ppm biphenyl as an internal standard, then shaking by Strong shaker SR-2 DS (TAITEC) at 300 rpm for 30 min. The extracted samples were left at room temperature overnight. 1 µl from an upper layer of the solution was used for gas chromatography (GC/FID) analysis.

2.2 Results

2.2.1 Cloning of GK2771 gene and construction of gene expression plasmid

GK2771

The *E. coli* DH5α contained pUC19 – GK2771 was obtained but due to unsuccessful cloning of pET28a - GK2771 by ligation into *Nco*I and *Eco*RI sites, a new insert was amplified aiming to ligate by blunt end instead. Only transformant from primer F71+atg (additional ATG next to the start codon) was obtained, Fig.18 showed result of colony PCR, there were many samples with a band near 900 bp indicating an existence of the insert. As a result of blunt end ligation, there were 2 possible directions for ligated insert, *Eco*RI enzyme was used to differentiate the direction (Fig. 19). From the gel, three types of patterns appeared, the highest band approximately 6000 bp was the correct one. From screening step, only two colonies that contained desired plasmid were found. The plasmids were sequenced and transferred into *E. coli* BL21 (DE3) for protein expression.

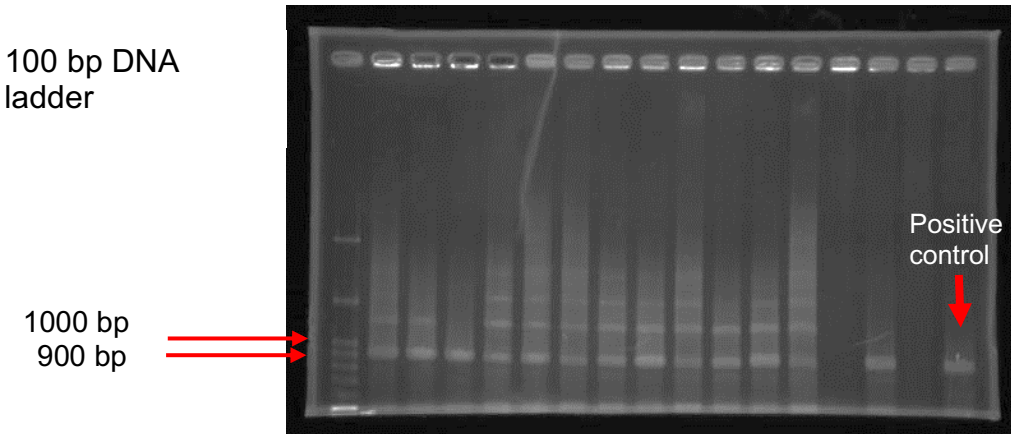


Fig. 18. Colony PCR of colonies from pUC19 - GK2771 transformation. Many samples showed the insert band at 909 bp.

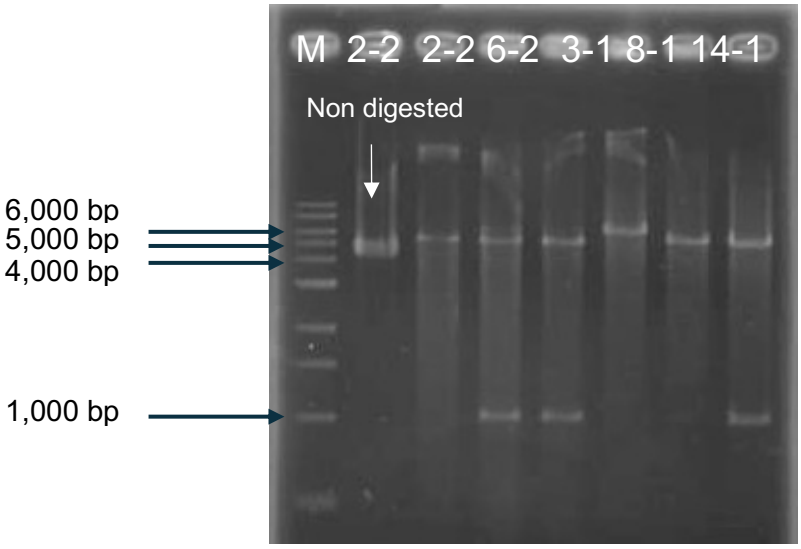


Fig. 19. Plasmid digestion by *Eco*RI enzyme, 3 types of patterns were observed, 1) 1 band below 6000 bp is a pET28a without insert, 2) 2 bands of 5,000 bp band and 1,000 bp band is a pET28a-GK2771 but undesired direction and 3) 1 band around 6000 bp is a pET28a-GK2771 with desired direction. Only 1 sample (8-1) contained a pET28a-GK2771 with desired direction.

Mutated GK2771

Mutated pET28a – GK2771 was prepared by KOD -Plus- Mutagenesis Kit which used inverse PCR technique, then the amplified plasmid was transformed to *E. coli* DH5α. Only 3 transformants were obtained, then the bacteria were cultured for plasmid extraction. Fig. showed the plasmid, 2 of the bacteria contained the plasmid which had similar size to pET28a – GK2771 (6,278 bp). The mutation was confirmed by sequencing, then the plasmid was transferred into *E. coli* BL21 (DE3) for protein expression.



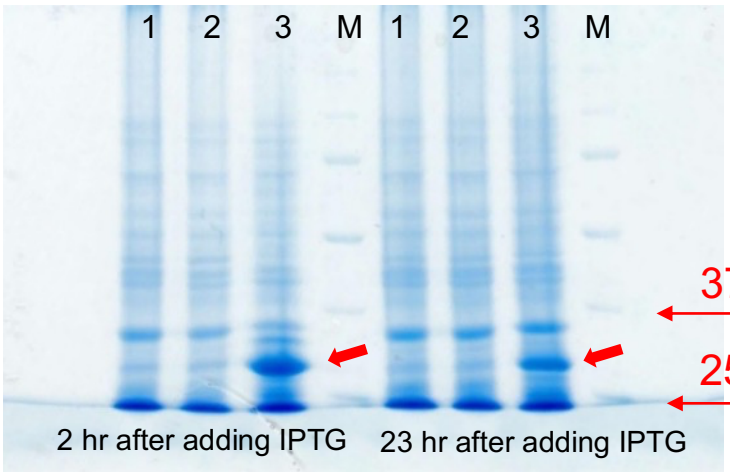
Fig. 20. Plasmid comparison between pET28a, pET28a – GK2771 and plasmids extracted from *E. coli* DH5α transformed with mutated pET28a – GK2771.

GK2771 with His tag

A new GK2771 fragment containing *Nde*I and *Eco*RI restriction site was amplified and ligated to *sma*I site of pUC19 without difficulty. The fragment was cloned into *Nde*I – *Eco*RI site of pET28a vector to utilize Histidine tag sequence located before the *Nde*I site. After some duration the transformant contained pET28a – GK2771 His tag transformants were successfully obtained, the plasmid was sequenced and transferred to *E. coli* BL21 (DE3). Subsequently, the mutated version of pET28a – GK2771 His tag was constructed by KOD - Plus- Mutagenesis Kit, using pET28a – GK2771 His tag as a template.

2.2.2 Heterologous gene expression of GK2771 and its variations in *E. coli*

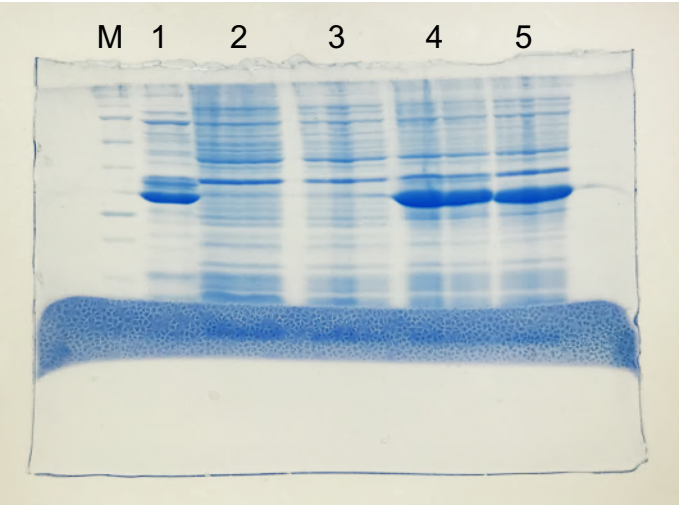
The protein production profiles of *E. coli* BL21 (DE3) pET28a - GK2771 after IPTG induction was showed in Fig. 21. A dense band appeared between 28 kD and 37 kD marker, clearly showing a difference between induced and non-induced sample. Although the expected protein molecular weight was 35.33 kD which calculated from the gene sequence the appeared band estimated size was around 30 kD (Molecular weight estimation, Bio-rad). Nevertheless, the recombinant protein was successfully expressed. In addition, protein expression of both plasmids obtained from ligation showed similar protein size.



M: Precision Plus Protein™ Unstained Protein Standards
1: *E. coli* BL21 (DE3) pET28a, not induced
2: *E. coli* BL21 (DE3) pET28a - GK2771, not induced
3: *E. coli* BL21 (DE3) pET28a - GK2771, IPTG induced

Fig. 21. The 10% SDS-PAGE of *E. coli* BL21 (DE3) pET28a-GK2771. The bands of a recombinant protein clearly appeared in both time conditions in the IPTG induced *E. coli* BL21 (DE3) pET28a-GK2771 culture.

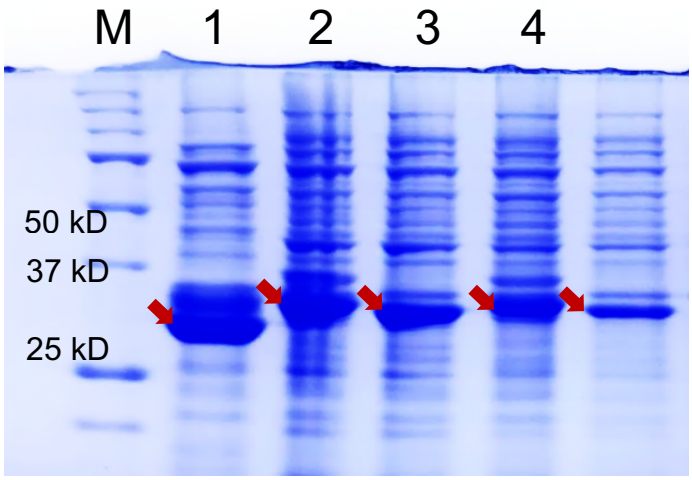
For the mutated GK2771, a similar protein band pattern to wildtype GK2771 was observed (Fig.22). The molecular weight of the mutated was also less than the calculated molecular weight.



M: Precision Plus Protein™ Unstained Protein Standards
1: Positive control, crude GK2771 Cell lysate of
2,3: *E. coli* BL21 (DE3) pET28a
4: *E. coli* BL21 (DE3) pET28a - GK2771
5: *E. coli* BL21 (DE3) pET28a - mutated GK2771

Fig. 22. The 15% SDS-PAGE of *E. coli* BL21 (DE3) cell lysate induced by 0.1 mM IPTG contained pET28a - GK2771 and pET28a - mutated GK2771.

For the his tagged protein, the calculated are as follows, GK2771 His tag is 37.36 kD and mutated GK2771 His tag is 37.24 kD. Both of his tagged proteins are slightly larger than the wildtype as expected but still smaller than the calculated molecular weight (Fig.23).



M: Precision Plus Protein™
Unstained Protein Standards
Soluble fraction of cell lysate
1: pET28a - GK2771
2,4: pET28a - GK2771 His tag
3,5: pET28a – mutated GK2771 His tag

Fig. 23. The 15% SDS-PAGE of *E. coli* BL21 (DE3) cell lysate induced by 0.1 mM IPTG, comparison between pET28a – GK2771 and his tagged variants.

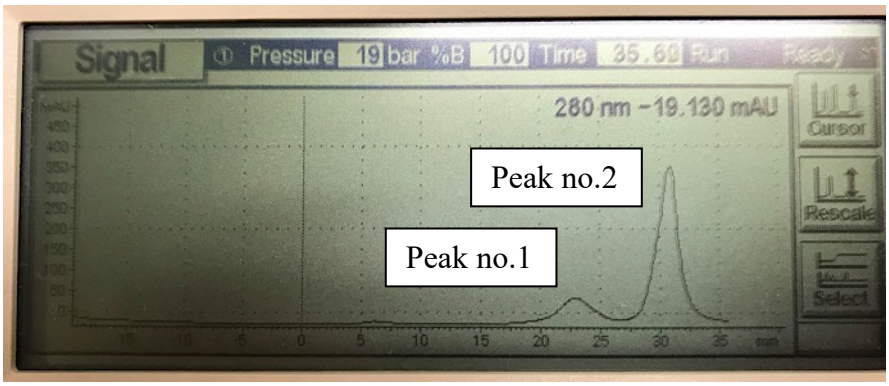
2.2.3 Protein purification

GK2771 and mutated GK2771

Anion exchange chromatography

The soluble fraction of *E. coli* BL21 (DE3) cell lysate was applied to HiTrap Q FF column, there was 2 peaks appeared within the elution gradient (Fig. 24a). The SDS PAGE revealed that the targeted protein was eluted within the first peak. However, the peak was not sharp, only high molecular weight proteins were removed from the targeted protein fraction (Fig. 25). Applying higher concentration of the cell lysate caused other peaks to appear during the purification but the peak which targeted protein exist also became clear (Fig. 24a).

a



b

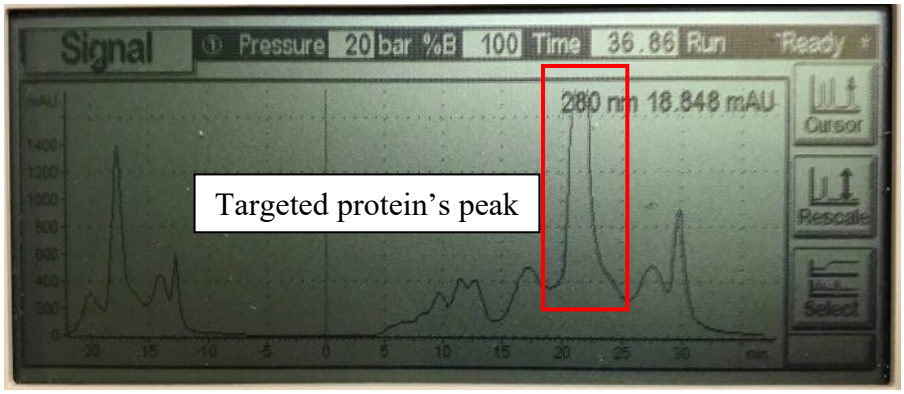


Fig. 24. The chromatogram from HPLC displays a signal during the anion exchange chromatography purification of *E. coli* BL21 (DE3) pET28a – GK2771 cell lysate. a) Low concentration sample, b) High concentration sample: Peak no. 1 becomes much larger, verifying that it is the targeted protein's peak.

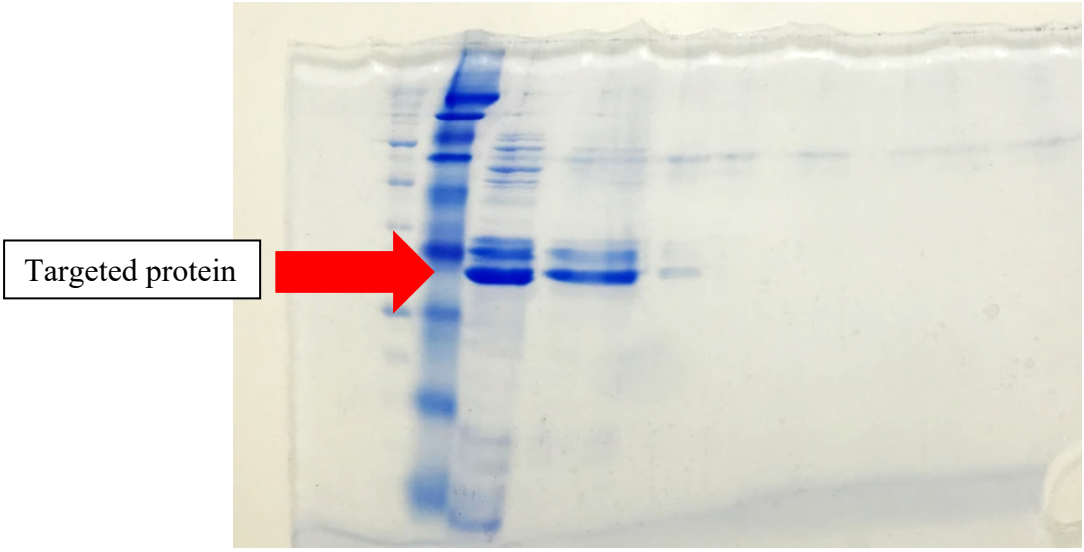


Fig. 25. The 15% SDS-PAGE of peak no.1 and no.2 eluted from HiTrap Q FF column, an anion exchange chromatography column. Eluted fraction from peak no.1 contained many proteins.

Hydrophobic Interaction Chromatography

All the protein from soluble fraction of *E. coli* BL21 (DE3) cell lysate was eluted out before entering the elution gradient (Fig. 26), that means the column and condition could not be used to purify this protein.

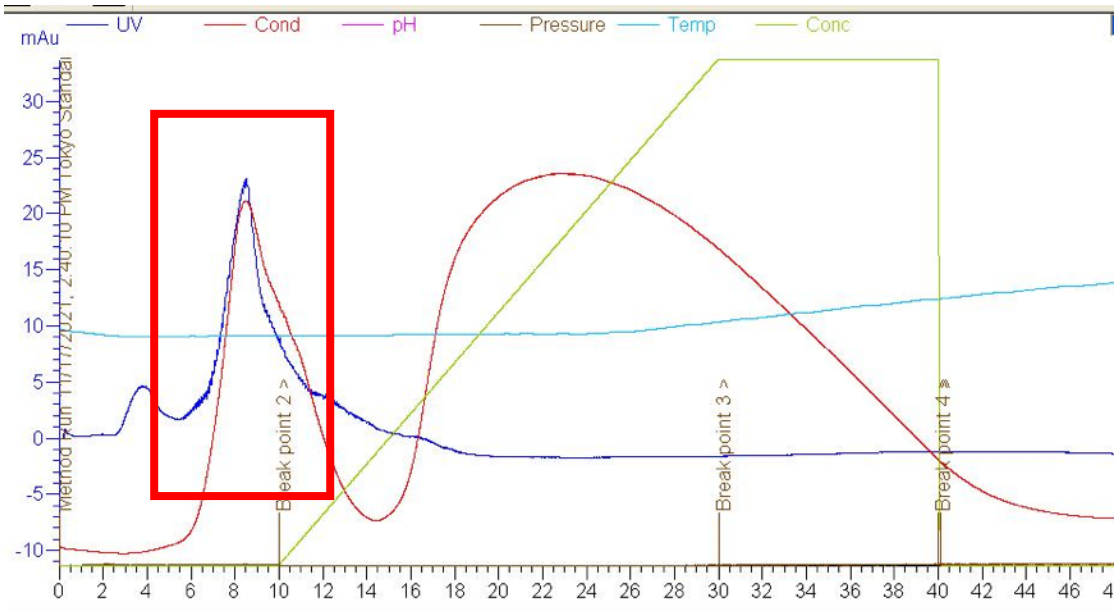


Fig. 26. The chromatogram from AKTA Prime Liquid Chromatography System displays a signal during the hydrophobic interaction chromatography purification of *E. coli* BL21 (DE3) pET28a – GK2771 cell lysate. All of the protein eluted before entering the gradient.

His tag purification, immobilized metal affinity chromatography

Imidazole used in elution of his tagged protein also has absorbance at 280 nm which we used to observe a protein peak during purification. The imidazole was supposed to have very low absorbance but the peak of proteins during trial purification with low concentration was masked by imidazole (Fig. 27), this led to a bit confusion. Anyway, SDS PAGE analysis of eluted fraction showed the GK2771 his tag position (~70 – 80%B) within the elution gradient. A small curve can be seen at the position and when high concentration of sample was applied to the column, the peak became visible. Unexpectedly, the purified GK2771 his tag contained 2 protein bands which N-terminal sequence analysis revealed that both proteins have histidine tag. The purified mutated GK2771 was also in the same condition (Fig. 28).



Fig. 27. The chromatogram from HPLC displays a signal during the His tag purification of *E. coli* BL21 (DE3) pET28a – GK2771 His tag cell lysate. The red circle highlights the part where the targeted protein signal was detected. The line appeared during 15 – 40 mins was caused by imidazole.

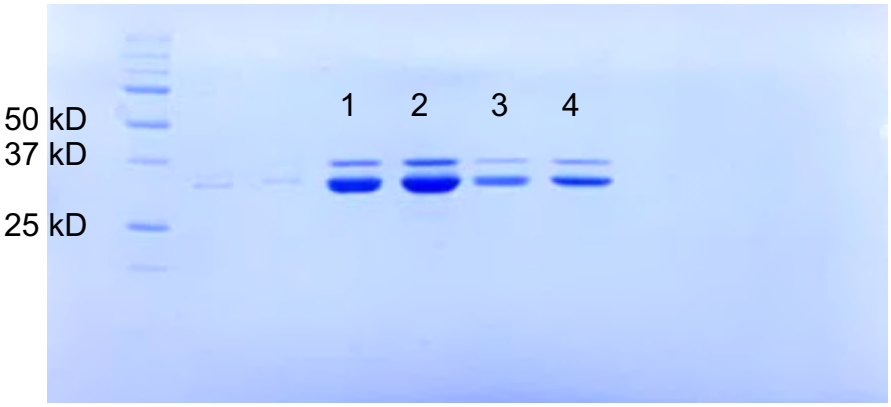


Fig. 28. The 15% SDS-PAGE of purified His tagged protein sample. 1: Acetone precipitated purified GK2771 his tag; 2: Acetone precipitated purified mutated GK2771; 3: Acetone precipitated purified GK2771 eluted from column; 4: Acetone precipitated purified mutated GK2771 eluted from column

2.2.4 Alkane degradation assay

Crude GK2771 and crude mutated GK2771

Alkane degradation by crude enzyme was examined at 37℃ and 60℃. The degradation was not significantly different between the two temperatures after considering the effect of evaporation on alkane at 60℃. The remaining alkane amounts after treatment with crude GK2771 for 24 h were reduced to around 70 – 80% except for the C10 alkane which was 38% compared to the negative control (autoclaved cell lysate) (Fig. 29).

Activity of the crude GK2771 and mutated GK2771 at 37℃ is shown in Fig.30, the mutated GK2771 was still able to degrade C10 – C17, although the difference in activity was not highly significant compared to the GK2771, the effect of mutation on activity was noticeable.

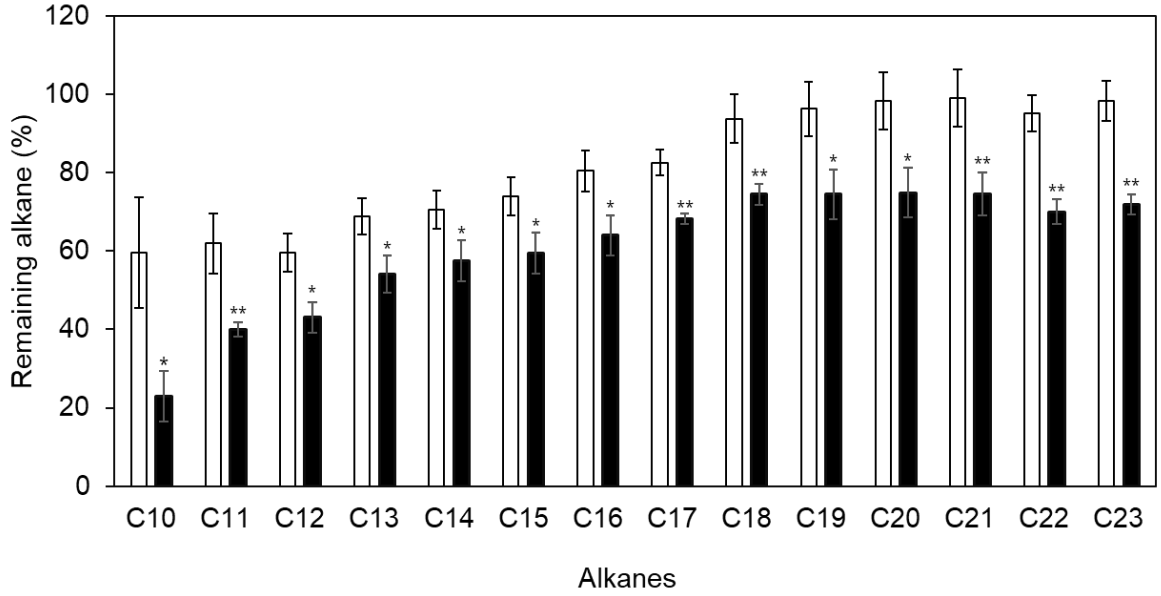


Fig. 29. Alkane degradation activity of crude GK2771 at 37 °C for 24 h. Autoclaved cell lysates were used as negative controls. White bar, autoclaved *E. coli* cell lysate; Black bar, GK2771-producing *E. coli* cell lysate. Single asterisk: $P < 0.05$, Double asterisk: $P < 0.01$

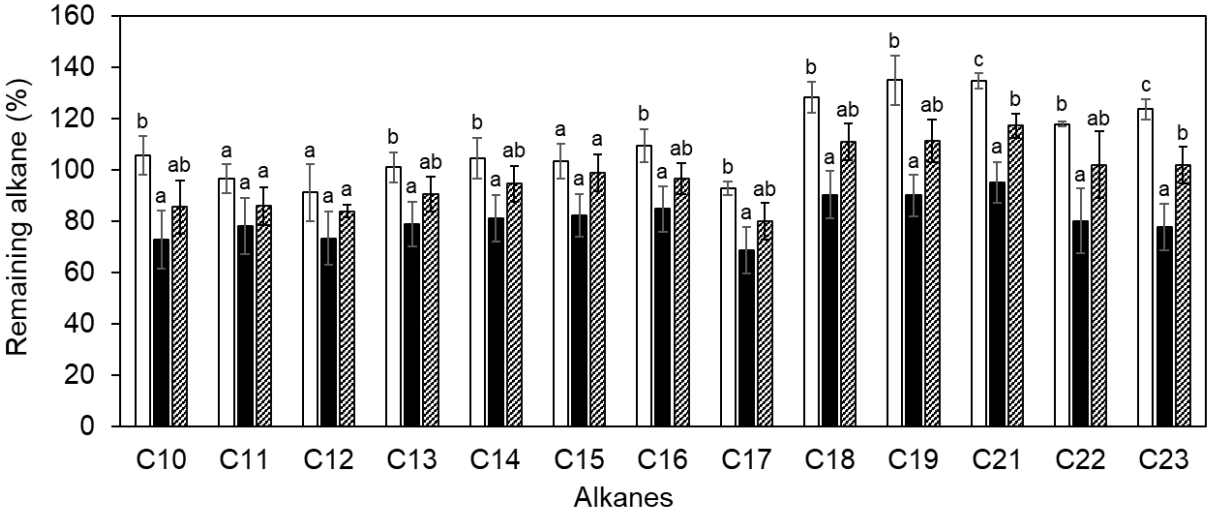


Fig. 30. Alkane degradation activity of crude GK2771 and crude mutated GK2771 at 37 °C for 24 h. Empty vector *E. coli* cell lysates were used as negative controls. White bar, *E. coli* pET28a cell lysate; Black bar, GK2771-producing *E. coli* cell lysate; Hatched bar, mutated GK2771-producing *E. coli* cell lysate. The letters indicate statistically significant ($P < 0.05$) differences.

Purified GK2771 His tag

Alkane degradation of crude GK2771 his tag was performed at 37℃ (data not shown) to confirm that the addition of histidine tag to GK2771 would not interfere with its alkane degradation activity before trying the alkane degradation with the purified GK2771 His tag.

At 37℃, 0.5 mg/ml of purified GK2771 His tag and the mutated counterpart did not display any degradation activity, the remaining alkanes were above 100% compared to the blank (Fig. 31a).

At 60℃, the purified GK2771 His tag could alkane degradation was observed. The C11 was reduced to 32.63% and other alkanes (C12 – C23) were reduced to around 70 – 80% compared to blank samples. The C10 was not detected in 2 out of 3 purified GK2771 samples.

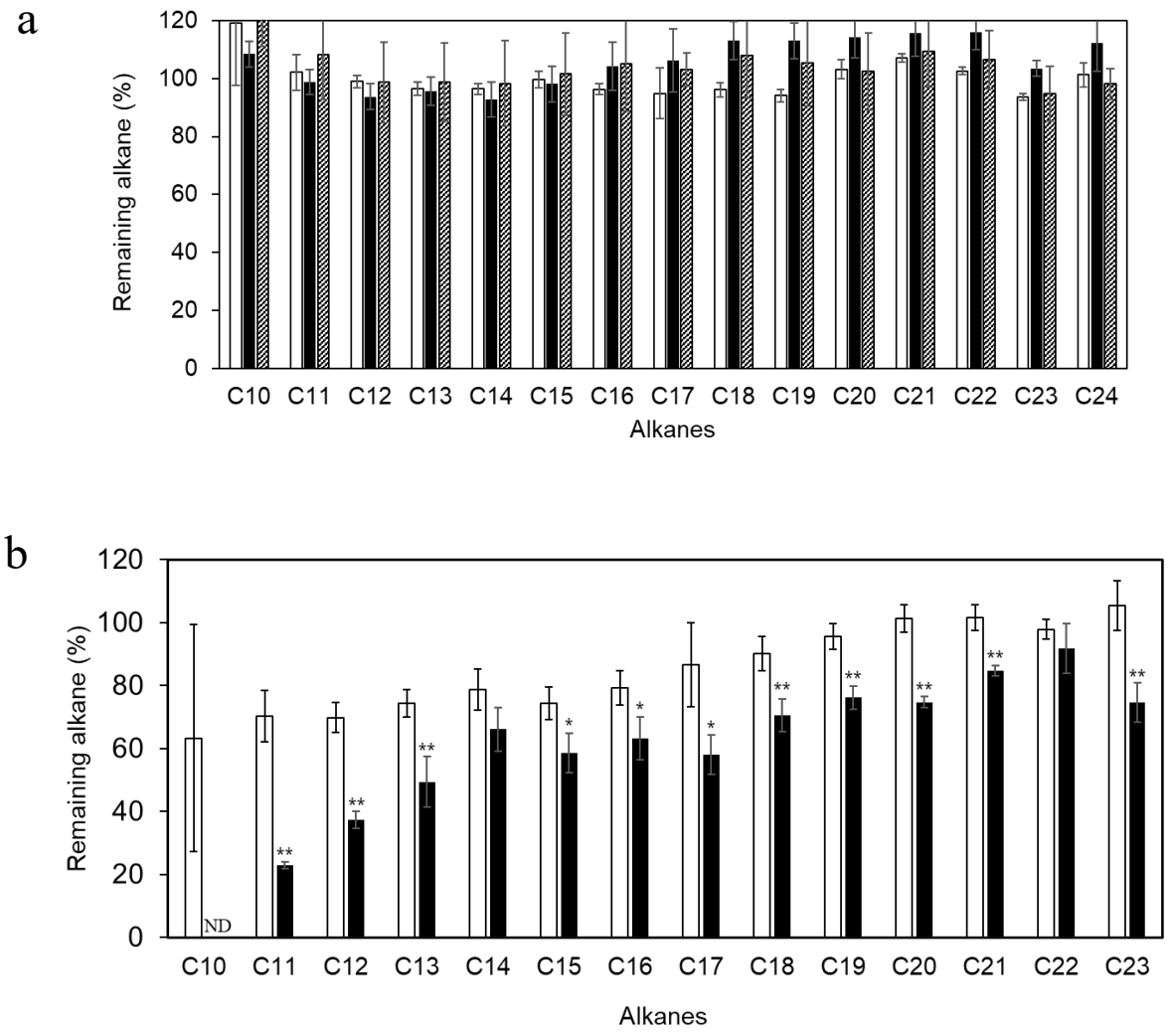


Fig. 31. Alkane degradation activity of purified GK2771 His tag and crude mutated GK2771 at a) 37 °C b) 60 °C for 24 h. Empty vector *E. coli* cell lysates were used as negative controls. White bar, buffer; Black bar, purified GK2771 His tag; Hatched bar, purified mutated GK2771. Single asterisk: $P < 0.05$, Double asterisk: $P < 0.01$

2.3 Discussion

According to the *G. kaustophilus* HTA426, the original bacteria alkane degradation ability, the GK2771 could degrade all type (C10 – C23) of the available alkanes is foreseeable. The difference between the enzyme and the bacteria degradation pattern is that the enzyme degraded a fair amount of C10 while the *G. kaustophilus* HTA426 seems to not prefer C10.

The mutation (E202A/ H205A) on GK2771 was aimed to stop the enzyme alkane degradation activity by changing two metal ligand amino acid residues (out of eight E69, V72, E102, H105, Y162, E167, E202, and H205) which attached to the Fe atom of the protein active site to alanine (Fig. 32). The result indicates that the two point mutations were not enough possibly due to the remaining six ligand amino acid residues still bound and functioning. Further mutation on other metal ligand could be performed to investigate the protein activity.

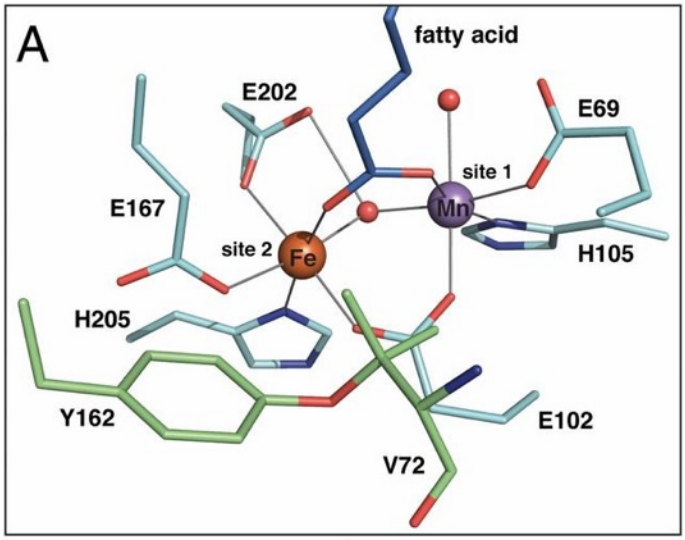


Fig. 32. Crystal structures of *Gk*R2loxI show metal ligand amino acids. (Griese et al., 2013) Reprinted with permission from Griese et al., 2013. Copyright 2024 American Chemical Society

Why crude enzyme could function at 37℃ but not the purified enzyme, the answer could be that the cell lysate contained necessary co-factor or enzyme catalyst for the enzyme although it was from *E. coli* not *G. kaustophilus* HTA426. While at 60℃, the increasing of temperature could gave enough energy for the enzyme to become active.

CHAPTER III

Heterologous gene expression of GK2772 gene in *E. coli*

3.1 Materials and methods

3.1.1 Strains and plasmids

Similar to GK2771, the heterologous expression of the GK2772 gene, has been done in *E. coli* BL21 (DE3) using pET28a and included the N terminal histidine tag. The bacterial strains and plasmids used or constructed in this study are listed in **Table 6.**

Table 6. Strains and plasmids used in heterologous gene expression of GK2772 gene (RNR2) in *E. coli* BL21 (DE3).

Strain or plasmid		Reference
Strains		
E. coli DH5a	Host strain for pUC19 cloning vector	-
E. coli BL21 (DE3)	F^- *ompT* $hsdS_B(r_B^- m_B^-)$ *gal dcm* (DE3) Host strain for pET gene expression vector Phage lamdaDE3 lysogenic strain containing T7 RNA polymerase gene under the control of lac promoter operator	pET system manual, Novagen
Cloning and expression vectors		
pUC19	Cloning vector, containing Ap^r gene	Yanisch-Perron *et al.,* 1985
pET28a	T7 promoter/ lac operator, containing Kan^r gene and lacI	pET system manual, Novagen
pUC19-GK2772 His	pUC19 with GK2772 gene with histidine tag	This study
pET28a-GK2772 His	pET28a with GK2772 gene with histidine tag	This study

3.1.2 Cloning of GK2772 gene and construction of gene expression plasmid

In the same manner to GK2771 gene cloning, the GK2772 gene was amplified by PCR with KOD-Plus-Neo DNA polymerase (Toyobo, Kyoto, Japan) from *G. kaustophilus* HTA426 genomic DNA. The primers, F72H (*Nde*I) and R72H (*Eco*RI) used for PCR amplification were designed by attaching *Nde*I and *Eco*RI restriction sites to the forward start and the reverse end of GK2771 gene sequence, respectively. The PCR amplified GK2772 gene with blunt end was cloned into *Sma*I site of pUC19. The heat shock method (Inoue et al., 1990) was used for *E. coli* DH5α transformation. The insert direction and no sequence error were confirmed by restriction enzyme digestion and sequencing, respectively. Next, the GK2772 gene was cut and cloned into *Nde*I and *Eco*RI restriction sites of pET28a vector without difficulty. The sequence with no error was confirmed and subsequently transferred into *E. coli* BL21 (DE3).

Table 7. Primers used in pET28a - GK2772 histidine tag construction

Genes	Length (bp)	Primer name	Sequences	Restriction site
gk2772	1497	F72His	5′ GACTGCATATGGTCCATCATGACGG 3′	*Nde*I
		R72His	5′ CGGAATTCTCATGACTCAGCAGCC 3′	*Eco*RI

3.1.3 Heterologous gene expression of GK2772 in *E. coli*

E. coli BL21 (DE3) harboring plasmid contained pET28a - GK2772 His tag was pre-culture in small volume of LB at 37℃, 120 rpm overnight. The culture was transferred to main culture by 1% inoculum size and continued incubation under the same condition until reaching mid log phase or an OD_{600} is 0.5-0.6 (approximately 2-3 hour). Subsequently, IPTG was added to final concentration 0.1 mM, and the culture was further incubated for an additional 4 hours. The whole culture was collected and centrifuged at 13,000 rpm for 10 min 4˚C, followed by removal of the supernatant. The cell pellet was resuspended in an appropriate buffer for further experiment and small volume aiming for high concentration of the protein. The resuspended pellet was sonicated using the Handy Sonic model CR-20P (TOMY Seiko co., ltd, Tokyo, Japan) on ice to prevent protein denaturation (4 cycles of 30 seconds sonication followed by 30 seconds resting period). The sonicated samples were centrifuged at 13,000 rpm, 10 min at

4˚C to discard the cell pellet. A portion of the sample was analyzed by SDS-PAGE to confirm the protein production. *E. coli* BL21 (DE3) contained the pET28a vector without insert and non-induced culture were used as negative controls.

3.1.4 Protein purification

Histidine tagged GK2772 purification

The purification process was similar to 2.1.4.2 which purification condition are as follows:

Column: HisTrap HP 1 ml (Cytiva)

Binding buffer (Buffer A): 20 mM sodium phosphate, 0.5 M NaCl, 40 mM imidazole, pH 7.4

Elution buffer (Buffer B): 20 mM sodium phosphate, 0.5 M NaCl, 500 mM imidazole, pH 7.4

The elution gradient was 15 min. After the protein's peak was identified, the elution method was changed from gradient to step wise elution. The gradient was replaced by 10 min each of following buffer B concentration 0%, 70%, 100% and 0% for total 40 min. The eluted protein fractions underwent desalting by PD-10 Desalting column (Cytiva) to replace the elution buffer with 0.1 M Tris buffer, pH 7.0. The protein concentration was measured by Bio-rad Protein Assay (Bio-rad) and the protein bands were analyzed on 15% SDS-PAGE.

3.1.5 Aldehyde dehydrogenase assay

GK2772's activity was tested against Acetaldehyde (C2), Decanal (C10) and Tetraldecanal (C14). Commercial yeast aldehyde dehydrogenase (MP Biomedicals) was used as a control. The preparation is as follows, dissolves 20 mg of the yeast ALDH lyophilizate in 1 ml of sterile water, then dilute 1:200 in ice cold 0.02% albumin solution.

Following Okibe et al., 1999 for soluble aldehyde, ALDH activity was determined by NADH production from the reaction which exhibits absorbance at 340 nm. The assay mixture (1 ml) contained 50 mM tricine buffer (pH 8.5), 1 mM NAD^+, 100 mM KCI, 2% Triton X-100, 1 mM acetaldehyde, and an appropriate amount of the enzyme (1 µg). The reaction was observed for 5 min at various temperature (30℃ - 70℃). At higher temperature the mixture

will become clouded possibly due to Triton X-100 interfering with the absorbance measurement.

For insoluble aldehyde the activity could not be measured by spectrophotometer. The activity was assessed by gas chromatography, directly measuring the reduction of amount substrate aldehyde. Reaction mixture was prepared in two different methods as follows:

Experiment 1: The assay mixture (0.5 ml) was prepared in 5 ml screw capped glass tube. In case of decanal 0.5 µl were added directly to mixture per sample, while tetradecanal will be mixed with buffer (2.12 mg/ 10 ml) and need to be heated for tetradecanal powder to totally dissolve (final concentration of substrate is 1 mM). A 0.5 ml mixture of 50 mM tricine buffer (pH 8.5) with 2% Triton X-100 and 100 mM KCl was added. Then 10 µl each of 50 mM NAD and enzyme were added. The mixture was briefly vortex and incubated in heat block to start enzyme reaction. After reaching the incubation time for 0, 10 and 60 min, 0.5 ml 6 M HCl and 2 ml methanol was added to stop the reaction. The mixture was incubated at 100℃ for 3 hours performing methanolysis of fatty acid. Each reaction was done in independent three bottles (N=3).

Experiment 2: The assay mixture (1.0 ml) was prepared in 5 ml screw capped glass tube.
A 1 ml mixture of 50 mM tricine buffer (pH 8.5) and 100 mM KCl was first added to the tube. Next, 20 µl each of 50 mM NAD^+ and 50 mM dodecanal or 50 mM tetradecanal dissolved in Triton X-100 were added to the tube and pre-heated for 5 min at various temperatures (30℃ - 70℃) using a heat block. After adding 10 µl of an enzyme, the mixture was briefly vortex and incubated in heat block to start enzyme reaction. After reaching the incubation time for 0, 10 and 60 min, added 50 µl of 6N HCl to stop the reaction. Each reaction was done in two bottles (N=2).

The whole mixture was extracted with equal volume (3 ml or 1 ml) of hexane containing 20 ppm dodecane as an internal standard using Strong shaker SR-2 DS (TAITEC) at 300 spm for 30 min. The extracted samples were centrifuged at 6000 rpm for 10 min, 2 µl from an upper hexane layer of the solution was used for gas chromatography (GC/FID) analysis. The GC system was same as the alkane degradation experiment.

3.2 Results

3.2.1 Cloning of GK2772 gene and construction of expression plasmid

The transformants contained pET28a - GK2772 His tag was successfully obtained. The plasmid was extracted and checked by double digestion (*Nde*I and *Eco*RI) and sequencing. Fig.33 shows 2 bands appeared at the expected size, pET28a is 5,369 bp and GK2772 is 1,497 bp. The sequencing result confirmed that there is no error in the gene sequence. The plasmid was transformed to *E. coli* BL21 (DE3) for protein expression.

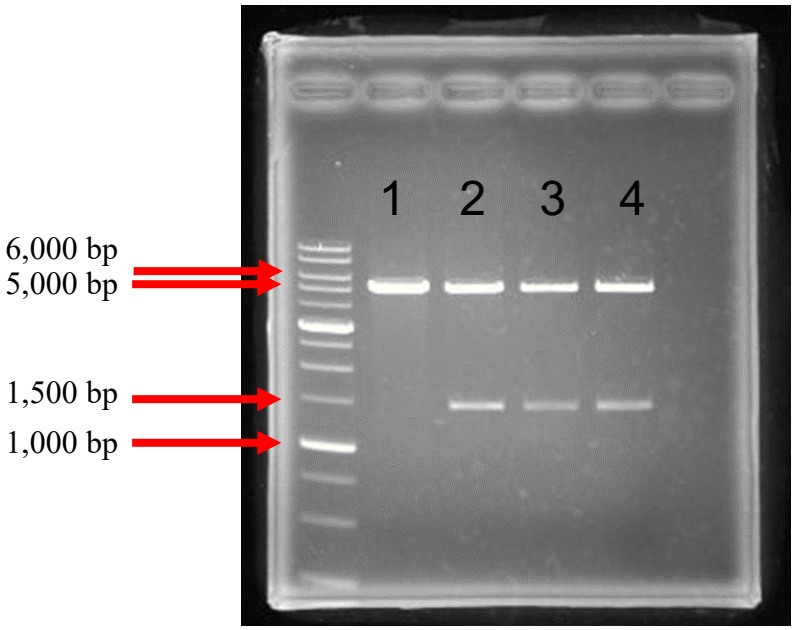


Fig. 33. Agarose gel displayed pET28a – GK2772 His tag digestion by *Nco*I and *Eco*RI. Two bands of pET28a (5,369 bp) and the insert DNA (1,497 bp) were clearly shown. Lane no.1: empty pET28a; Lane no.2 – 4: plasmid extracted from the selected pET28a – GK2772 His tag transformants

3.2.2 Heterologous gene expression of GK2772 in *E. coli*

The protein production profiles of *E. coli* BL21 (DE3) pET28a - GK2772 His tag after IPTG induction was shown in Fig. 34. A dense band appeared above the 50 kD marker band, which clearly showed a difference between induced and non-induced sample. The recombinant protein band estimated size is in line with the calculated protein molecular weight which is 56.02 kD.

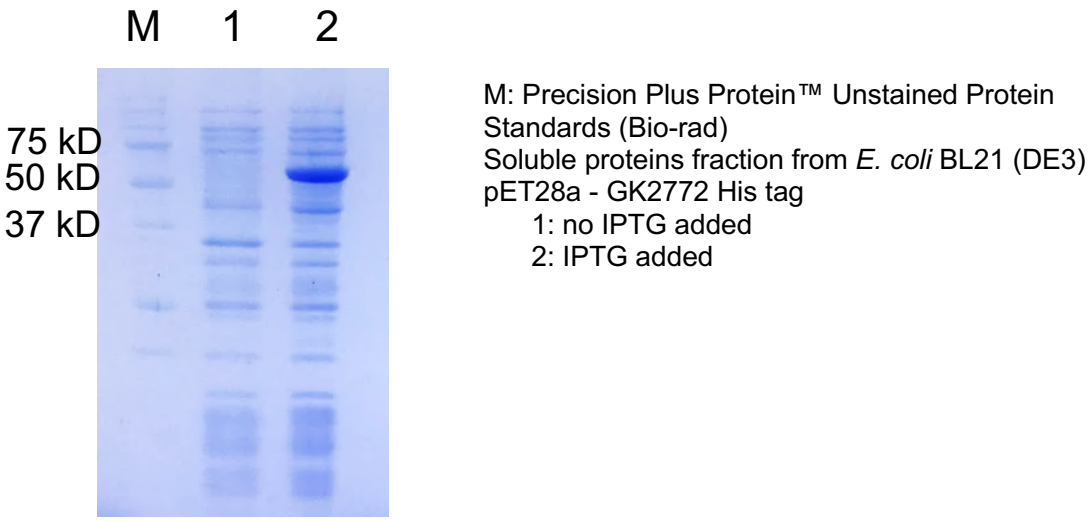
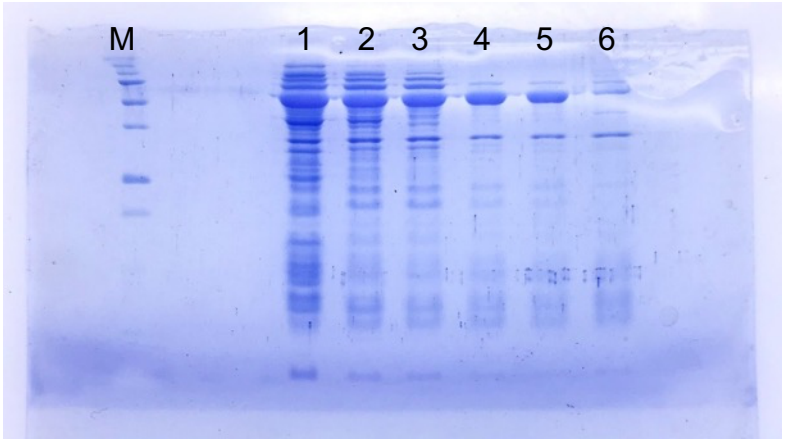


Fig. 34. The 15% SDS-PAGE of *E. coli* BL21 (DE3) pET28a - GK2772 His tag. The bands of a recombinant protein clearly appeared in the IPTG induced culture.

3.2.3 Protein purification

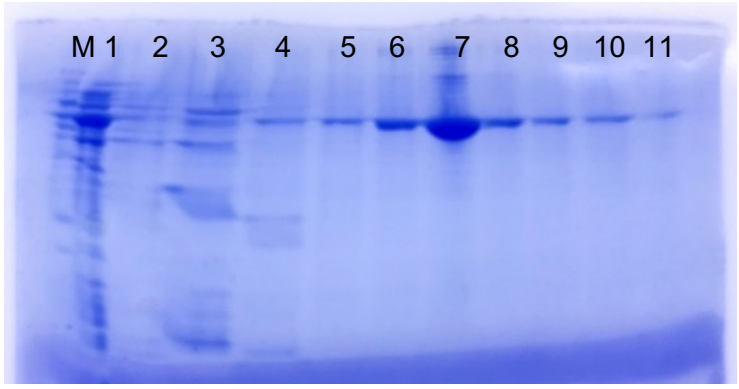
The protein sample was heat treated at 70℃, followed by centrifugation to remove denatured protein (Fig. 35), prior to application onto the His trap column. At first, under gradient elution condition, the targeted protein eluted out across various %B concentration (Fig. 35a). After several adjustments, the protein was eluted as single peak at 70%B (Fig. 36). However, the eluted fraction still contained other protein in small quantity.



M: Precision Plus Protein™ Unstained Protein Standards (Bio-rad)
1: not treated
2: 50℃
3: 60℃
4: 70℃
5: 80℃
6: 90℃

Fig. 35. The 15% SDS-PAGE of *E. coli* BL21 (DE3) pET28a - GK2772 His tag cell lysate soluble fraction after 10 min heat treatment.

a



M: Precision Plus Protein™ Unstained Protein Standards (Bio-rad)
1: the protein sample before purification
eluted fractions
2: 14.2 – 15.7 mins
3: 16.0 – 17.5 mins
4: 19.4 – 20.9 mins
5: 20.9 – 22.4 mins
6: 22.4 – 23.9 mins
7: 23.9 – 25.4 mins
8: 25.4 – 26.9 mins
9: 26.9 – 28.4 mins
10: 28.4 – 29.9 mins
11: 29.9 – 31.4 mins

b

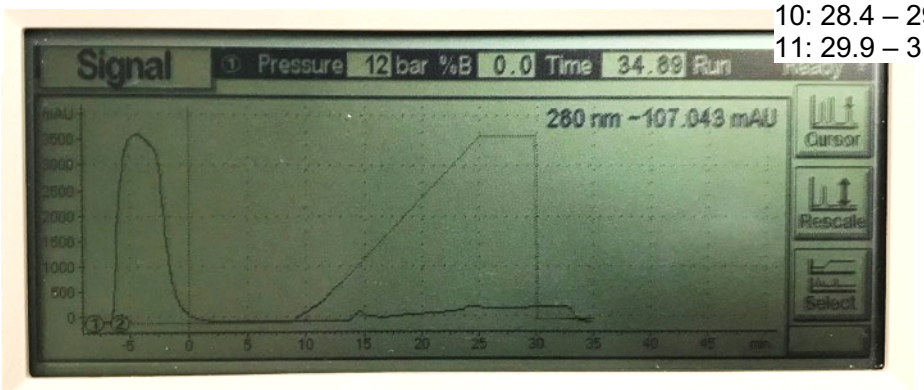


Fig. 36. a) The 15% SDS-PAGE of GK2772 His tag protein fraction eluted from column chromatography purification. At this step many fractions contain part of targeted protein. b) The chromatogram from HPLC displays a signal during the His tagged protein purification.

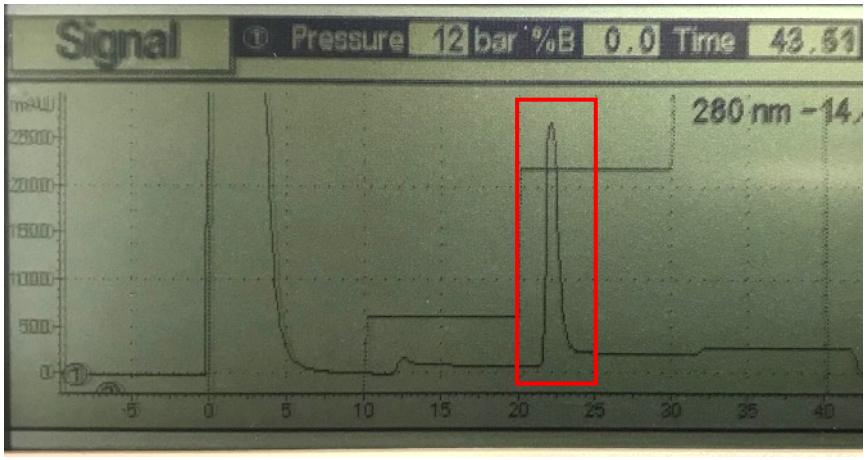


Fig. 37. The chromatogram from HPLC displays a signal during the His tagged protein purification. High concentration sample was used in this purification, the peak of targeted protein was clearly shown at 70%B.

3.2.4 Aldehyde dehydrogenase assay

By spectrometer

Against 1 mM acetaldehyde, over 5 min at 37℃ increasing of NADH absorbance (340 nm) was observed for yeast ALDH but not purified GK2772. At 60℃ both enzymes showed activity. Next, the activity of GK2772 was tested at different temperatures (30℃, 40℃, 50℃ and 60℃), except at 30℃ which enzyme was not active. GK2772 showed the highest initial rate and activity at 60℃. While at 40℃ and 50℃, the activity was almost the same. When the reaction was further tested above 60℃, the mixture became clouded, and the absorbance could not be measured.

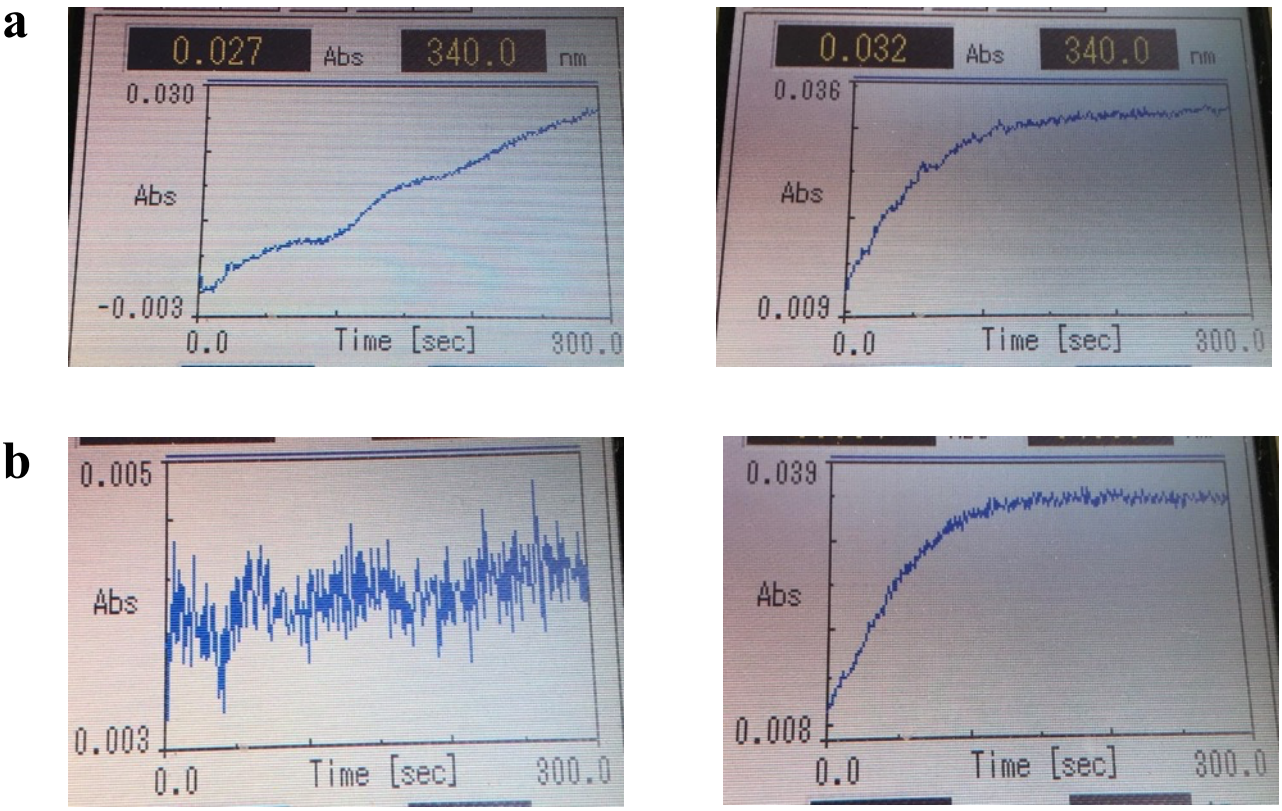


Fig. 38. Absorbance at 340 nm from aldehyde dehydrogenase assay from a) commercial yeast ALDH and b) purified GK2772 over 5 min. Left: at 37℃; Right: 60℃

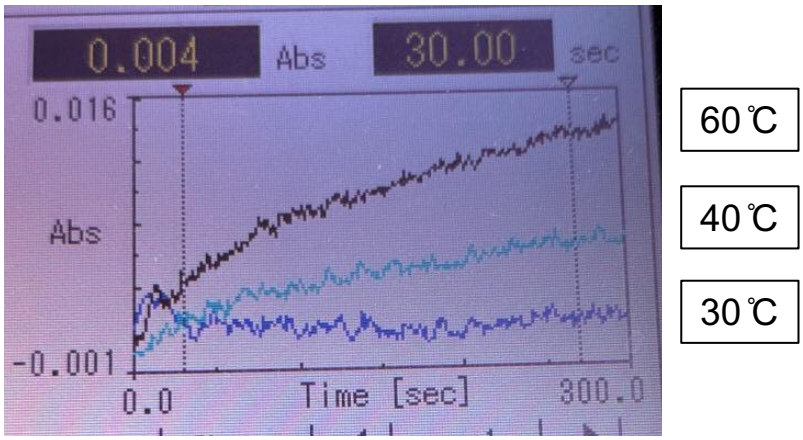
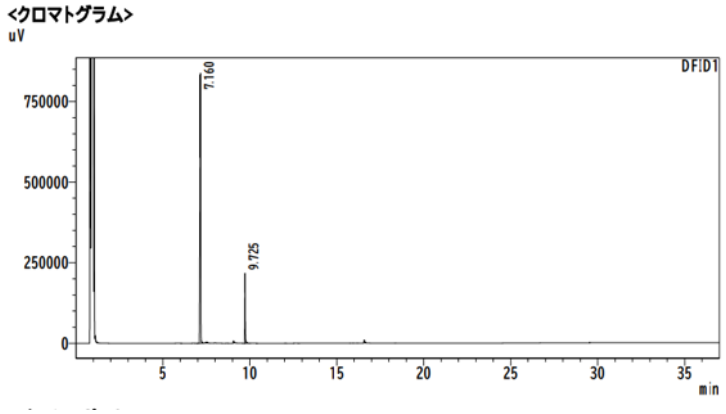


Fig. 39. Combined result of absorbance at 340 nm from aldehyde dehydrogenase assay from purified GK2772 over 5 min. Black: 60℃; Cyan: 40℃; Blue: 30℃

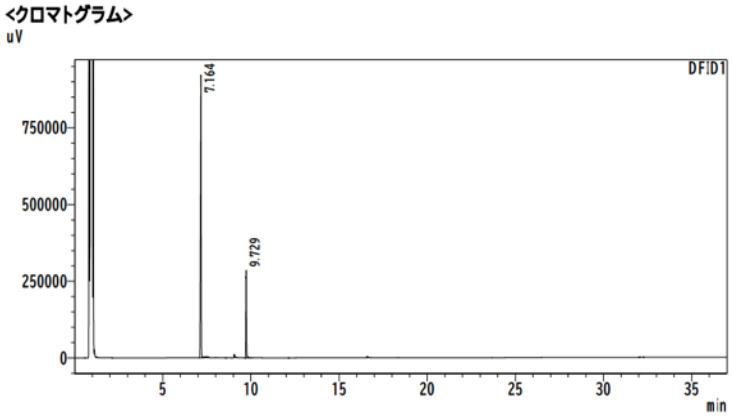
By gas chromatography

Initial attempt for the gas chromatography analysis indicated no aldehyde dehydrogenase activity in both samples at 60℃. There was no difference between blank, yeast ALDH and GK2772 samples. Neither a decrease of decanal nor appearance of product peak was observed following 10-minute and 60-minute incubation periods. Except for single GK2772 sample among triplicate, each per incubation time. The reduction of decanal and a new peak of assumed to be a reaction product were noticed. (Fig. 40c and 40e) The remaining amount of decanal remained at 100% or above, while the one that reaction occurred, the decanal remained 29.19% (10 min) and 21.45% (60 min). The experiment using tetradecanal as a substrate also did not show any activity.

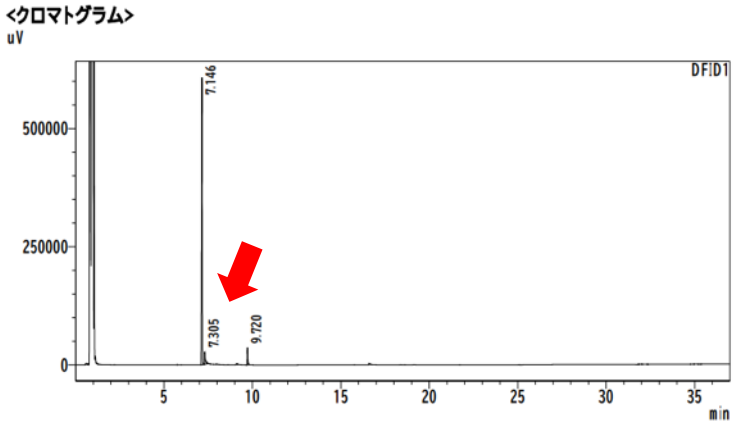
a



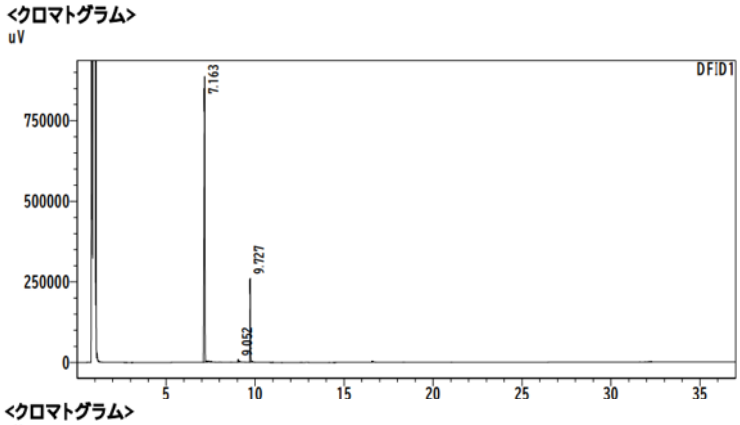
b



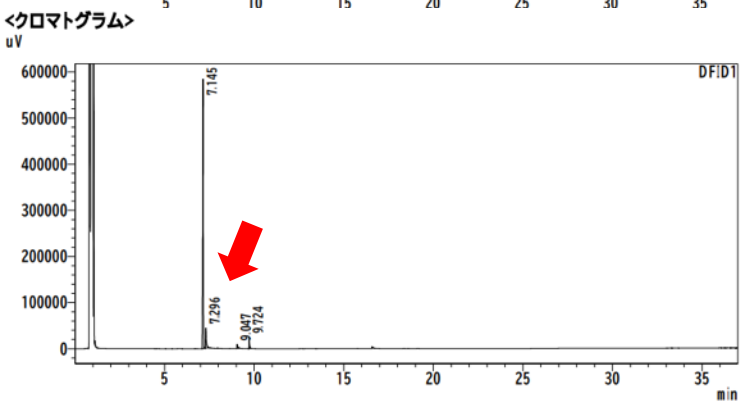
c



d



e



f

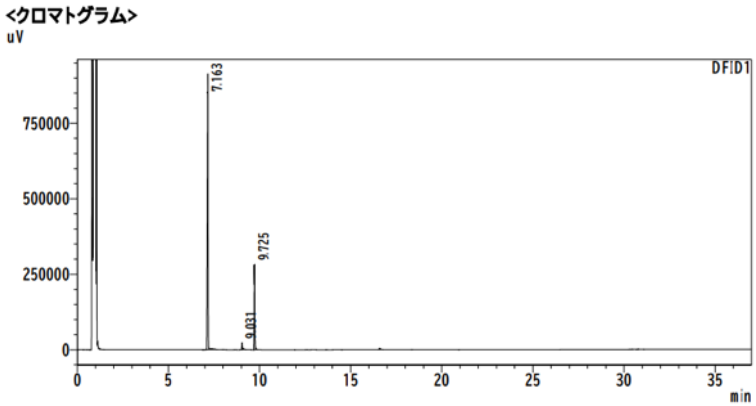
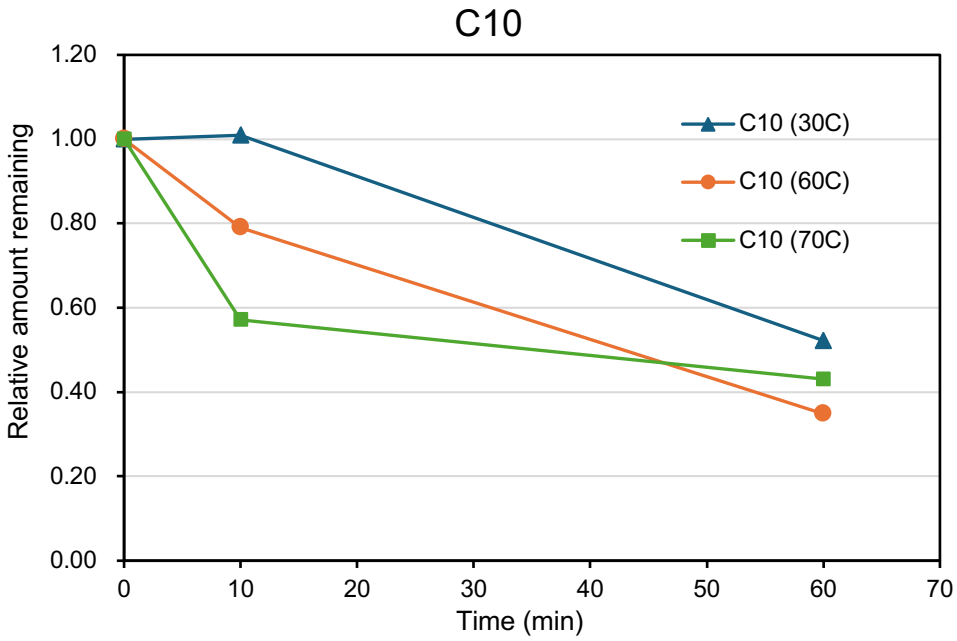


Fig. 40. Gas chromatography chromatogram result of aldehyde dehydrogenase assay at 60℃ (1st experiment, underwent methanolysis) a) Blank, decanal 0 min; b) Blank, decanal 10 min; c) GK2772 protein, decanal 10 min bottle no.1, d) GK2772 protein, decanal 10 min bottle no.2; e) GK2772 protein, decanal 60 min bottle no.2; f) GK2772 protein, decanal 60 min bottle no.3.

Peaks at 7.1 min and 9.7 min are of dodecane, internal standard, and decanal, substrate, respectively. The red arrow is assumed to be a peak of reaction product decanoic acid methylester. The samples (c,e) which reaction occurred the amount of decanal was clearly reduced (~9.7 min peak) compared to other sample.

The aldehyde dehydrogenase assay was retried with modification, where the substrate was directly dissolved in Triton X-100 expected to increase the solubility. The methanolysis step was omitted. In this experiment, the reduction of both decanal and tetradecanal was clearly observed in the sample containing GK2772 whereas in the blank samples aldehyde remained as initial amount during incubation (Fig. 41). For decanal, the reaction was tested at 30˚C, 60˚C and 70˚C, for 10 min and 60 min incubation periods. At 10 min incubation time, 70˚C sample showed the highest activity (57.2% remaining), while at 30˚C sample there was no reduction observed. At 60 min, 60˚C sample showed the highest activity instead (34.8% remaining). It was found that 30˚C sample also showed significant activity (52.2% remaining) at 60 min incubation time. For tetradecanal, the experiment was conducted only at 60˚C. Tetradecanal was also reduced to 89.4% at 10 min and 43.9% at 60 min.

a



b

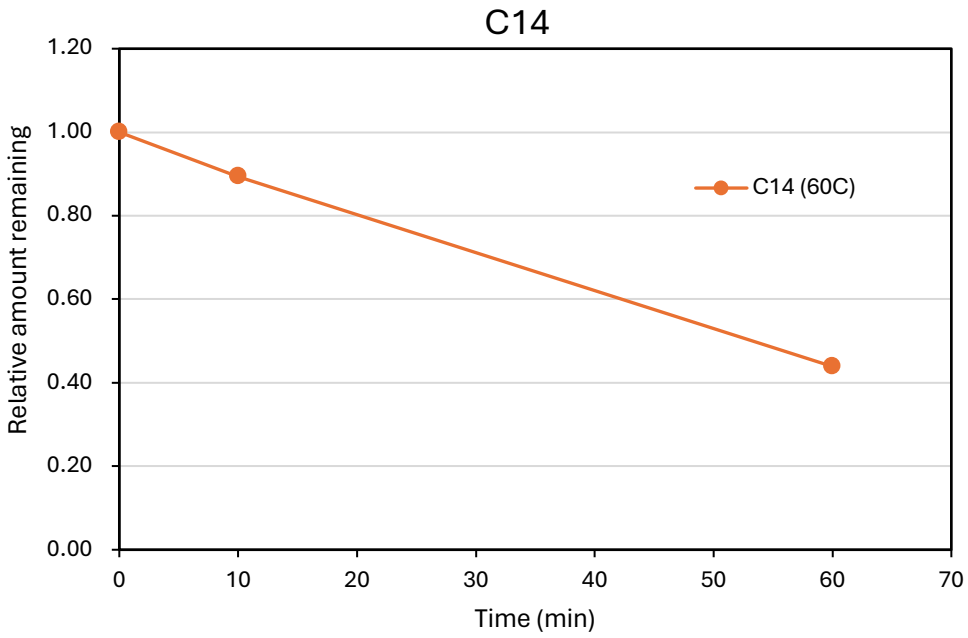
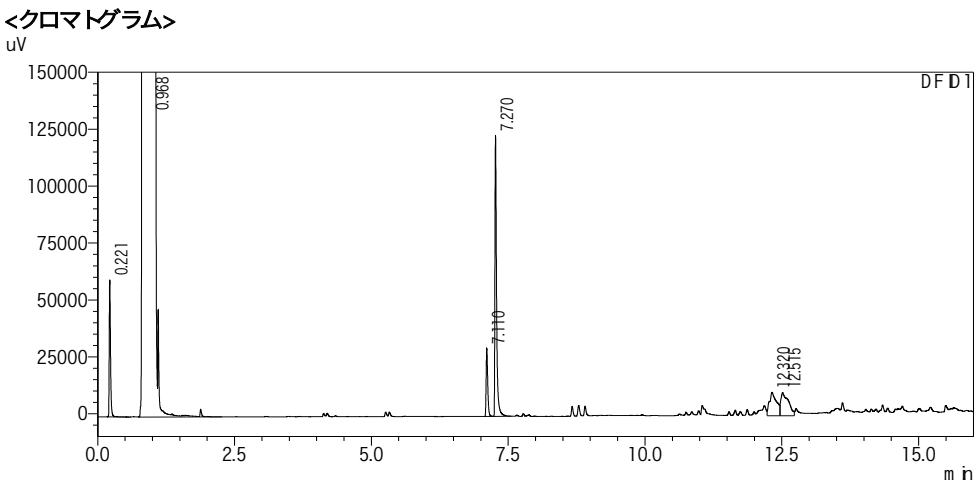


Fig. 41. Aldehyde dehydrogenase activity of purified GK2772 His tag (2nd experiment, did not undergo methanolysis) on a, decanal (C10) and b, tetradecanal (C14) for 10 min and 60 min incubation time. No enzyme added reaction sample was used as negative control. Symbols are triangle, 30˚C; circle, 60˚C; square, 70˚C. Data points are average of two samples (N=2).

a



b

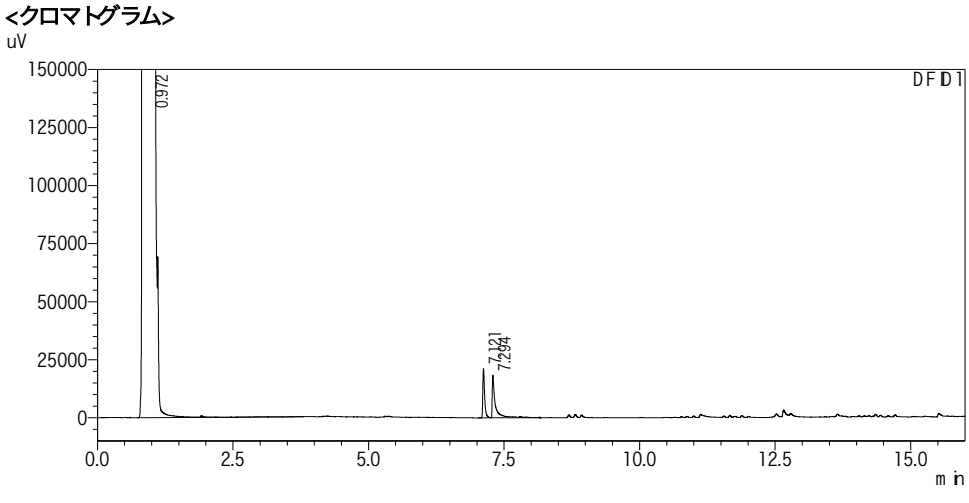


Fig. 42. Gas chromatography chromatogram result of aldehyde dehydrogenase assay at 60℃ (2nd experiment, did not undergo methanolysis) showing a reduction of amount decanal. a) 0 min ; b) 60 min. Peaks at 7.1 min and 7.2 min are of dodecane, internal standard, and decanal, substrate, respectively.

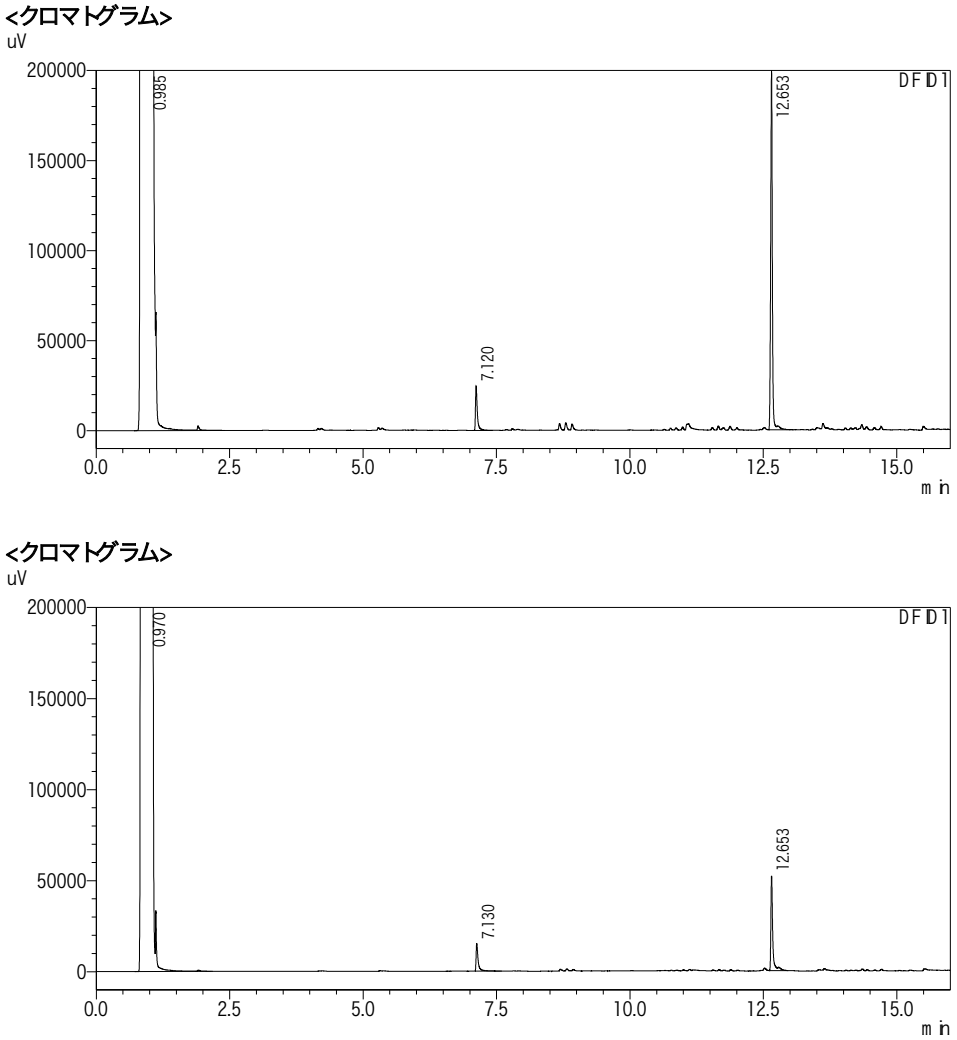


Fig. 43. Gas chrom,atography chromatogram result of aldehyde dehydrogenase assay at 60℃ (2^{nd} experiment, did not undergo methanolysis) showing a reduction of amount tetradecanal. a) 0 min ; b) 60 min. Peaks at 7.1 min and 12.6 min are of dodecane, internal standard, and tetradecanal, substrate, respectively.

3.3 Discussion

This experiment was done in order to obtain more evidence to support an idea that a set of GK2769-GK2782 gene cluster is responsible for alkane metabolisms. It was shown that the gene product of GK2772 was an aldehyde dehydrogenase which could react not only acetaldehyde (C2) but decanal (C10) and tetradecanal (C14) specifically at high temperature conditions, 60℃ and 70℃. This result is consistent with an observation that *G. kaustophilus* HTA426 degrades middle to long chain alkanes.

The aldehyde amount in blank samples was unchanged clearly showed that the reduction of aldehyde was caused by the enzyme (data not shown). The reason is unclear why the reaction did not stably occur in the first experiment, however, substrate aldehydes may not have been completely dissolved in 2% Triton X-100. The product's peak was undetected in the second experiment, possibly due to the fatty acid was not converted to fatty acid methylester form in the sample preparation step. Although the enzyme showed higher activity at higher temperatures for 10 min, the activity was decreased at 70˚C after 60 min incubation. This thermal instability may be reasoned by the fact that the enzyme was produced with an additional N-terminal His tag peptide. At 10-minute incubation time, 70˚C sample had lower remaining aldehyde amount but for 60-minute incubation time, the 60˚C sample had lower remaining aldehyde.

CHAPTER IV

General discussion

In this study, I demonstrated that GK2771 or *Gk*R2loxI (WP_011232245) is a novel alkane monooxygenase responsible for alkane degradation ability of *G. kaustophilus* HTA426 and successfully provided new knowledge on R2lox protein group. However, several topics require further investigation.

Alkane uptake and Enzyme specificity

Bacteria may acquire water insoluble alkanes through direct contact with alkane droplets or by producing biosurfactant to emulsify alkanes (Wang & Shao, 2013). Little information is available on substrate specificity of the alkane transport systems. From the results of *G. kaustophilus* HTA426 alkane degradation in LB, bacterium HTA426 prefers C11 – C17 alkanes more than others that are available in the experiment (Fig. 12). In contrast, HTA426 almost evenly degraded various alkanes, from C10 to C24 in LBM, where alkane was the sole carbon source. These results indicated that the capacity of the substrate alkane transportation system is not a determining factor for the preference of HTA426. This was probably due to its low binding affinity for the alkane monooxygenase/hydroxylate enzyme (data not shown). It is probable that the C11 – C17 alkanes could be the most energy efficient carbon source for HTA426; therefore, the degradation of this range of alkanes occurred even under nutrient rich conditions, LB.

A small amount of C9 alkane was observed in the standard gas oil used in this study, but reduction did not occur, and it is possible that the enzyme does not work or has a low affinity for alkanes smaller than C10. For an alkane longer than C24, further experiments need to be conducted to determine the size of the long chain alkane as the capacity limit of this enzyme. Because the activity of enzyme on C23 and C24 was low, I expected that the activity on longer chain alkane would not be much different.

The amounts of each alkane in the standard oil used in the alkane degradation experiments were unequal (substrate availability). In the future, equal amounts of alkanes (representative from each group such as short chain alkane, C10; medium chain alkane, C16; and long chain alkane C20) will be used for alkane degradation experiments or comparison of cell growth on a single alkane could reveal or confirm the alkane preference of *G. kaustophilus* HTA426.

In addition to GK2771 substrate alkane specificity, GK2772 (ALDH) and GK2774 (ADH) substrate specificity should also influence the preference of *G. kaustophilus* HTA426.

Origin of *Geobacillus kaustophilus* HTA426

G. kaustophilus HTA426 was isolated from the deepest sea mud sample at 10,897 m (100 MPa, 4˚C) of the Mariana Trench. Its characteristics were described as a thermophile (growing temperatures 55 – 75˚C), and it could not grow under 30 MPa hydrostatic pressure (in the paper 30 MPa, 60 MPa and 100 MPa were tested). Most of the isolated thermophiles from the mud sample were unable to grow under this high pressure. This led to a question of whether *G. kaustophilus* HTA426 live and grow there because of the low temperature and high hydrostatic pressure conditions of the sea bottom. It is conceivable that the isolated thermophiles including the *G. kaustophilus* HTA426 were in a dormant state at the deepest sea mud (Takami et al., 1997).

Moreover, based on metagenome analysis of water samples from the bottom of the Mariana Trench. MAGs possess predicted pathways for AlkB and AlmA. They also analyzed n-alkane (C15 – C32) concentration of the sinking particles from 2,000 m (7.2 μg/gdw), 4,000 m (39.0 μg/gdw) and 6,000 m (24.2 μg/gdw), unfortunately they were unable to analyze sample from deeper depths. While surface sediments at ~10,910 m has total n-alkane concentration averaged 2.3 μg/gdw (Liu et al., 2019). This study suggests the presence of alkane and alkane degraders at the bottom of the Mariana trench. Whether this is related to the acquisition of *G. kaustophilus* HTA426's alkane degradation gene still needs further investigation.

Distribution of GK2771 gene homologues in other bacteria

A Protein BLAST (Basic Local Alignment Search Tool) search revealed that homologous genes of GK2771/*Gk*R2loxI can be found scattered within the Geobacillus genus (with identities ranging from 85% to 100% compared to GK2771), as well as in other Bacilli species (48% to 98%), and even beyond the Bacilli class (40% to 66%).

Exploring through NCBI (National Center for Biotechnology Information) genome database and KEGG (Kyoto Encyclopedia of Genes and Genomes) database to compare the bacteria that contain homologous genes of GK2771, several gene patterns were recognized, including mesophilic bacteria.

First pattern, the bacteria contain all R2lox, ALDH, ADH and beta oxidation related gene that has high similarity and gene cluster is arrange in the same order (similar to GK2771 – GK2783 and GK3008) to *G. kaustophilus* HTA426 including hypothetical protein.

E.g. *G. thermoleoverans* CCB_US3_UFS (GTCCBUS3UF5_31160, 100% identity to GK2771) while other gene %identity is around 99%. The bacterium was isolated from Ulu Slim hot spring in Malaysia (Sakaff et al., 2012).

Geobacillus sp. Y412MC52 (GYMC52_2807, 91.06% identity), other gene %identity is range from 91 -100%. It was obtained from obsidian hot spring in Yellowstone National Park, USA (Brumm et al., 2015).

G. stearothermophilus 10 (GT50_05355, 91.06% identity), other gene %identity is 90 - 99%. Also isolated from hot springs in Yellowstone National Park, USA.

G. thermoleoverans KCTC 3570 (GT3570_13555, 89.07%), other gene %identity is 79.44 – 100%.

Second pattern, the bacteria possess all R2lox, ALDH, ADH and beta oxidation related genes with high similarity, but the gene order are dissimilar, yet they remain closed within the same genomic region.

E. g. *G. subterraneus* KCTC 3922 (GS3922_02395, 89.07% identity), other gene %identity is 88 - 98%. It was isolated from high-temperature formation waters of the Liaohe oilfield in China (Lee et al., 2017).

Parageobacillus caldoxylosityticus ER4B (BV455_02563, 78.63%), a whole gene cluster is presented (65 – 80% identity). It was previously known as *Geobacillus caldoxylosilyticus*, isolated from an empty oil palm fruit bunch compost in Malaysia (Ching et al.,2022).

Third pattern, the bacteria have R2lox, ALDH and ADH without beta oxidation gene cluster. Some related genes may be present, but they are not homologous to *G. kaustophilus* HTA426's gene.

E.g. *Anoxybacillus geothermalis* NRS-147 (P9711_17305, 93.95%), R2lox is accompanied by high similarity ALDH (P9711_17300, 92.91% to GK2772) and partial ADH (P9711_17290, 48% coverage, 93.89% to GK2774) Two acyl-CoA dehydrogenase gene (P9711_17385 and P9711_17390) also presented but have no similarity.

Caenibacillus caldisaponilyticus B157 (CCB157_RS00135, 69.05%), it contains distant ALDH (CCB157_RS00065,33.74% to GK2772) and ADH (CCB157_RS00070, 30.90% to GK2774).

Thermaerobacter composti ASM3358203v1(Q5761_RS07755, 40.14%), accompanied by ALDH (Q5761_RS07750, 51.25% to GK2772) and ADH (Q5761_RS07760, 31.09% to GK2774). It was obtained from geothermal fluids, Insheim, Germany.

Fourth pattern, the mesophilic bacteria have R2lox and some part of the cluster.

E.g. *Neobacillus cucumis* WH12 (QNK12_11935, 12% coverage, 70.65%), although it has ALDH, ADH and some other beta oxidation related gene but overall, very low similarity and dissimilar gene location. It was obtained from rhizosphere soil, Anhui, China.

Caldibacillus thermoamylovorans SSBM (CQJ30_09330, 68.03%), it possesses ALDH (no similarity to GK2772) and some part of the beta oxidation related gene cluster (66 – 75% identity). It was isolated from biodrying product from sewage sludge treatment plant, Shanghai, China (Cai et al., 2018).

Bacillota bacterium isolate SMAG_U10532 (VK136_06170, 63.67%), it has equivalent of GK2777 – GK2783 (52 – 82.50% identity) except a long chain fatty acid CoA ligase gene (VK136_06175) is not locate in beta oxidation gene cluster and has low similarity to GK2782 of *G. kaustophilus* HTA426 (74% coverage, 27.59% identity). The genome is obtained from reconstructing 40,039 soil microbial metagenome-assembled genome bins (the SMAG catalogue) project (Ma et al.,2023).

Caldibacillus debilis GB1 (RK062275, 62.46%), ALDH and ADH are absent, but it contains equivalent of GK2777 – GK2782 and GK3008 (27 – 76 % identity). It was isolated from an air-tolerant cellulolytic consortium (Wushke et al.,2015).

N. muris DSM110989 R2lox2 (NFS43_RS11795, 47.19%) and R2lox3 (NFS43_RS12680, 27.94%). The bacteria contain 3 R2lox proteins. NFS43_RS11795 is accompanied by ALDH (NFS43_RS11800, 49.58% to GK2772) and ADH (NFS43_RS11805, no similarity to GK2774). NFS43_RS12680 has acyl-CoA dehydrogenase (NFS43_RS12670, 30.10% to GK2781), long chain fatty acid CoA ligase (NFS43_RS12665, 45.61% to GK2782) and unrelated acetyl CoA acyltransferase. It was isolated from mouse cecal content (Afrizal et al., 2022).

Thermus scotoductus SA-01 (TSC_C16720, 42.96%), beside R2lox it has ALDH gene (TSC_C16730,50.83% identity to GK2772). It was obtained from fissure water collected 3.2 km below surface in a South African gold mine (Gounder et al.,2011).

Fifth pattern, the mesophilic bacteria have only the R2lox without other gene resembles *G. kaustophilus* HTA426's gene cluster close by.

E.g. *N. muris* DSM110989 R2lox1 (NFS43_RS04160, 48.25%) (Afrizal et al., 2022)

Anaerolineae bacterium UBA8986 (DHW49_15150, 46.28%) (Parks et al.,2018)

Fimbriimonadales bacterium KatS3mg018 (KatS3mg018_2647, 42.52%), isolated from sediments collected from the Katase hot spring field, Shizuoka, Japan (Kato et al.,2022).

Chloroflexi bacterium CFX2 (FBQ99_17700, 45.23%), isolated from anammox reactor sludge, Beijing, China.

Candidatus Neomarinimicrobiota bacterium XB_MAG1229 (VX822_05270, 41.50%), isolated from marine water sample, South Pacific Ocean (Tully et al.,2018).

Armatimonadota bacterium HOT.MAX.070 (K6U12_08935, 41.30%), isolated from Lakelse Hot spring, Terrace, Canada (Levy-Booth et al.,2022).

Roseiflexacea bacterium ES-bin-105 (H7242_01350, 41.20%), isolated from the Little Firn glacier, Greenland.

Bacteria harboring R2lox are distributed across diverse parts of the earth and exhibit varying metabolic characteristics, including aerobes, facultative anaerobes, and strict anaerobes. Most of them are thermophiles isolated from hot springs with some being mesophile. Despite the structural resemblance of R2lox to RNR2 of class Ic, it is possible that bacteria utilizing RNR from other classes can possess R2lox. In particular, anaerobic or facultative anaerobic species that are incapable of utilizing class I RNR, besides possessing multiple classes of RNR, are not uncommon.

This suggests that the percentage identity to GK2771 and the degree of gene cluster conservation are directly correlated. As well as the relation of the genetic similarity to *G. kaustophilus* HTA426 and those that contain *Gk*R2loxI-like gene. Considering that the percentage identity to GK2771 tends to decrease among taxa furth er from Geobacillus, Bacillus, and then outside of the Bacilli class.

The presence of the entire gene cluster (R2lox, ALDH, ADH, and beta oxidation related genes) suggests and supports the fact that GK2771 functions as alkane degradation, with conservation of the cluster likely maintaining its enzymatic capability. Bacteria possessing the complete cluster are very likely to degrade alkanes, similar to *G. kaustophilus* HTA426. While the R2lox sequences exhibiting low similarity to GK2771 are less likely to be associated with alkane degradation and may have evolved independently or diverged from a common ancestor with GK2771. Alternatively, this phenomenon could represent a possibility where R2lox as an alkane degradation protein could be developed at a later stage in thermophilic bacillus species.

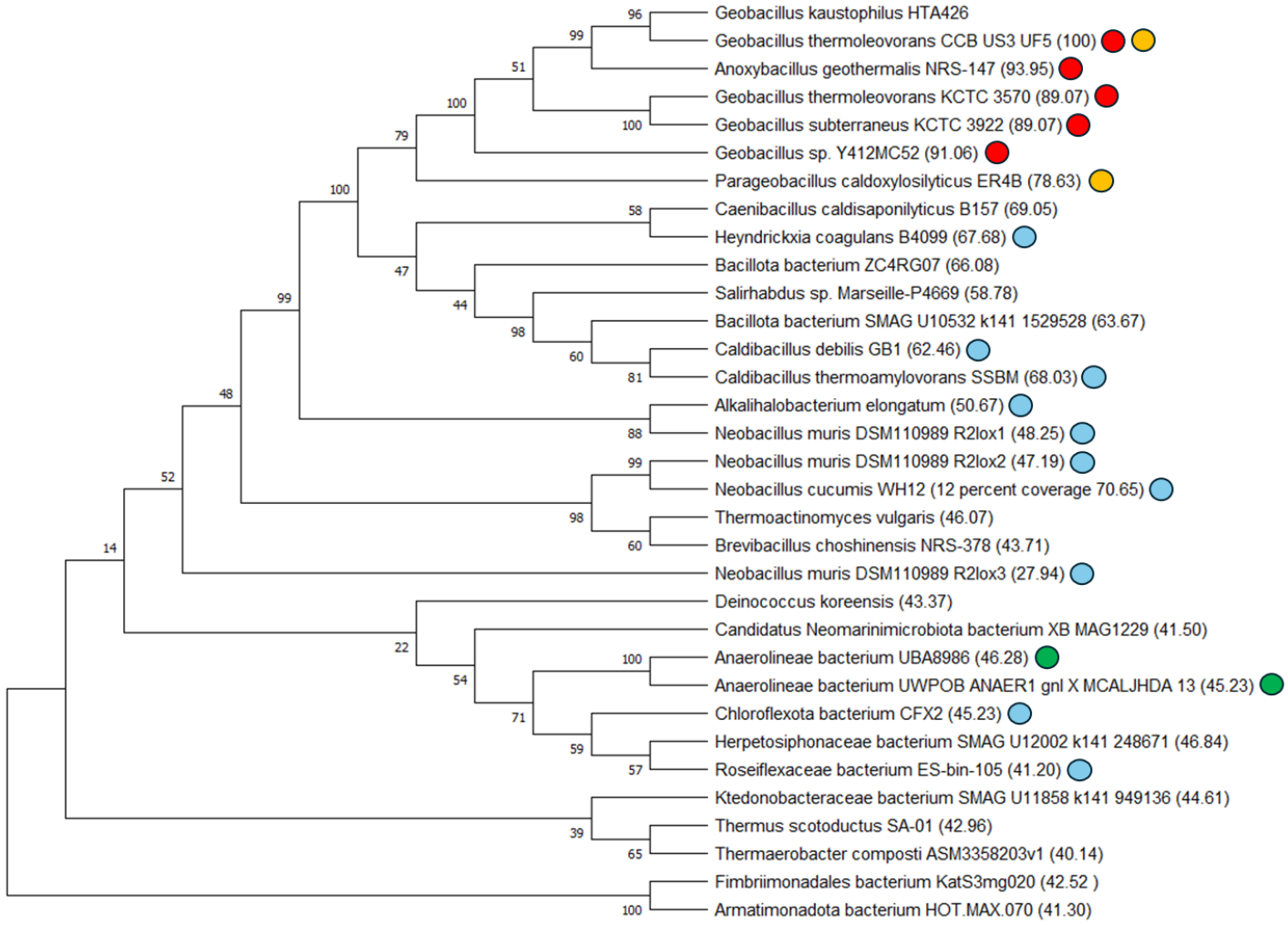


Fig. 44. A phylogenetic tree was constructed based on the protein sequences of R2lox from various bacterial species, comparing them to GkR2lox. The numbers in parentheses next to each bacterial species indicate the percentage identity to the GK2771 amino acid sequence. Evolutionary analysis was performed using the Maximum Likelihood method in MEGA11 with bootstrap samplings. 🔴: The bacteria contain whole GK2771 gene cluster; 🟡: The bacteria contain LadA type alkane monooxygenase; 🔵: Facultative anaerobe; 🟢: Strictly anaerobe

GKP26

G. kaustophilus HTA426 contains another R2lox gene candidate in its cryptic plasmid (pHTA426), called GKP26 or *Gk*R2loxII. Similar to GK2771, GKP26 was also identified as ribonucleotide reductase small subunit (RNR2) and also followed by alcohol dehydrogenase (GKP30) and aldehyde dehydrogenase (GKP31) but has no further step related gene in alkane metabolism. The identity between GK2771 and GKP26 amino acid sequences is only 30%. Whether GKP26 is involved in alkane or other hydrocarbon degradation needs to be further examined.

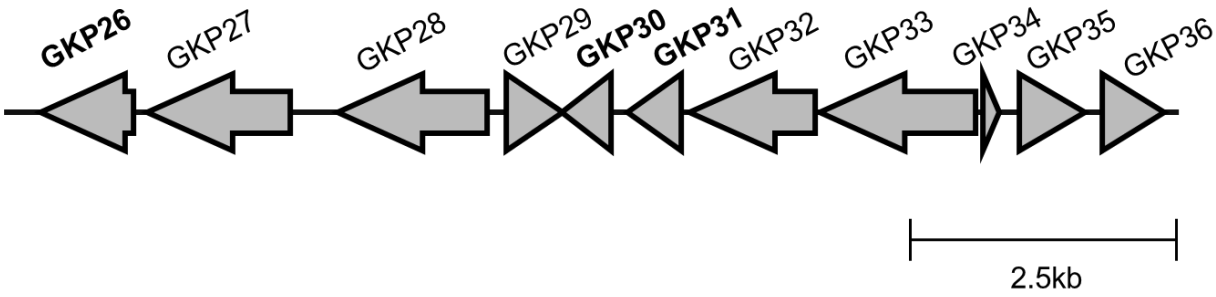


Fig. 45. A gene cluster GKP26 – GKP36 in the plasmid pHTA426 of *G. kaustophilus* HTA426. GKP26: ribonucleotide reductase small subunit, GKP27: MFS transporter, GKP28: 2,5-furandicarboxylate decarboxylase/ 4-hydroxybenzoate decarboxylase, GKP29: hypothetical protein, GKP30: alcohol dehydrogenase, GKP31: glycine betaine aldehyde dehydrogenase, GKP32: hypothetical protein, GKP33: hypothetical protein, GKP34: flavin prenyltransferase/ phenylacrylic acid decarboxylase, GKP35: hypothetical protein, GKP36: chromosome partitioning protein

ACKNOWLEGEMENT

First, I would like to express my deepest gratitude to my advisor Prof. Masaaki Morikawa for all the support, kindness, and patience leading me through the research and study.

To all the laboratory members, Kusakabe-san who help me with documents and many things, every past and present member you made me felt comfortable and happy to be here.

To friend in Sapporo and Japan that I met here and from Thailand, who made my stay enjoyable and content.

To family with unwavering support and never pressured me on what I do.

To MEXT Japan for awarding me scholarship during through Master and PhD course at Graduate School of Environmental Science, Hokkaido University.

It was a really long journey; I am really grateful for this experience.

REFERENCES

Afrizal, A., Jennings, S. A. V., Hitch, T. C. A., Riedel, T., Basic, M., Panyot, A., Treichel, N., Hager, F. T., Wong, E. O-Y., Wolter, B., Viehof, A., von Strempel, A., Eberl, C., Buhl, E. M., Abt, B., Bleich, A., Tolba, R., Blank, L. M., Navarre, W. W., Kiessling, F., Horz, H-P., Torow, N., Cerovic, V., Stecher, B., Strowig, T, Overmann, J., & Clavel, T. (2022). Enhanced cultured diversity of the mouse gut microbiota enables custom-made synthetic communities. *Cell Host & Microbe, 30*(11), 1630-1645.e1625. doi:10.1016/j.chom.2022.09.011

Alaine, B. G. Oliveria, J. S., Silva-Portela, R. C., Araújo,W., Carlos, A. C., Vasconcelos, A. T. R., Greitas, A. T., Domingos, Y. S., de Farias, M. F., Fernandes, G. J. T., & Angez-Lima, L. F. (2018). Metagenome enrichment approach used for selection of oil-degrading bacteria consortia for drill cutting residue bioremediation. *Environmental Pollution, 235*, 869-880. doi:https://doi.org/10.1016/j.envpol.2018.01.014

Andersson, C. S., & Högbom, M. (2009). A *Mycobacterium tuberculosis* ligand-binding Mn/Fe protein reveals a new cofactor in a remodeled R2-protein scaffold. *Proceedings of the National Academy of Sciences, 106*(14), 5633-5638. doi:10.1073/pnas.0812971106

Boll, M., Fuchs, G., & Heider, J. (2002). Anaerobic oxidation of aromatic compounds and hydrocarbons. *Current Opinion in Chemical Biology, 6*(5), 604-611. doi:https://doi.org/10.1016/S1367-5931(02)00375-7

Boonmak, C., Takahashi, Y., & Morikawa, M. (2014). Cloning and expression of three ladA-type alkane monooxygenase genes from an extremely thermophilic alkane-degrading bacterium Geobacillus thermoleovorans B23. *Extremophiles, 18*(3), 515-523. doi:10.1007/s00792-014-0636-y

Brumm, P., Land, M. L., Hauser, L. J., Jeffries, C. D., Chang, Y.-J., & Mead, D. A. (2015). Complete genome sequences of Geobacillus sp. Y412MC52, a xylan-degrading strain isolated from obsidian hot spring in Yellowstone National Park. *Standards in Genomic Sciences, 10*(1). doi:10.1186/s40793-015-0075-0

Cai, L., Zheng., S-W., Shen, Y-J., Zheng, G-D., Liu, H-T., & Wu, Z-Y.. (2018). Complete genome sequence provides insights into the biodrying-related microbial function of Bacillus thermoamylovorans isolated from sewage sludge biodrying material. *Bioresource Technology, 260*, 141-149. doi:https://doi.org/10.1016/j.biortech.2018.03.121

Ching, X. J., Najimudin, N., Cheah, Y. K., & Wong, C. (2022). Complete genome sequence data of tropical thermophilic bacterium Parageobacillus caldoxylosilyticus ER4B. *Data Brief, 40*, 107764. doi:10.1016/j.dib.2021.107764

Diamanti, R., Srinivas, V., Johansson, A. I., Nordström, A., Griese, J. J., Lebrette, H., & Högbom, M. (2022). Comparative structural analysis provides new insights into the function of R2-like ligand-binding oxidase. *FEBS Letters, 596*(12), 1600-1610. doi:10.1002/1873-3468.14319

Elumalai, P., Parthipan, P., Karthikeyan, O. P., & Rajasekar, A. (2017). Enzyme-mediated biodegradation of long-chain n-alkanes (C32 and C40) by thermophilic bacteria. *3 Biotech, 7*(2). doi:10.1007/s13205-017-0773-y

Feng, L., Wang, W., Cheng, J., Ren, Y., Zhao, G., Gao, C., Tang, Y., Liu, X., Han, W., Peng, X., Liu, R., & Wang, L. (2007). Genome and proteome of long-chain alkane degrading Geobacillus thermodenitrificans NG80-2 isolated from a deep-subsurface oil reservoir. *Proc Natl Acad Sci U S A, 104*(13), 5602-5607. doi:10.1073/pnas.0609650104

Gounder, K., Brzuszkiewicz, E., Liesegang, H., Wollherr, A., Daniel, R., Gottschalk, G., Reva, O., Kumwenda, B., Srivastava, M., Bricio, C., Berenguer, J., van Heerden, E., & Litthauer, D. (2011). Sequence of the hyperplastic genome of the naturally competent Thermus scotoductus SA-01. *BMC Genomics, 12*(1), 577. doi:10.1186/1471-2164-12-577

Griese, J. J., Roos, K., Cox, N., Shafaat, H. S., Branca, R. M. M., Lehtiö, J., Gräslund, A., Lubitz, W., Siegbahn, P. E. M., & Högbom, M. (2013). Direct observation of structurally encoded metal discrimination and ether bond formation in a heterodinuclear metalloprotein. *Proceedings of the National Academy of Sciences, 110*(43), 17189-17194. doi:10.1073/pnas.1304368110

Högbom, M., Stenmark, P., Voevodskaya, N., McClarty, G., Gräslund, A., & Nordlund, P. (2004). The radical site in chlamydial ribonucleotide reductase defines a new R2 subclass. *Science, 305*(5681), 245-248. doi:10.1126/science.1098419

Inoue, H., Nojima, H., & Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. *Gene, 96*(1), 23-28. doi:10.1016/0378-1119(90)90336-p

Ji, Y., Mao, G., Wang, Y., & Bartlam, M. (2013). Structural insights into diversity and n-alkane biodegradation mechanisms of alkane hydroxylases. *Front Microbiol, 4*, 58. doi:10.3389/fmicb.2013.00058

Jiang, W., Yun, D., Saleh, L., Barr, E. W., Xing, G., Hoffart, L. M., Maslak, M-A., Krebs, C., & Bollinger, J. M., Jr. (2007). A manganese(IV)/iron(III) cofactor in Chlamydia trachomatis ribonucleotide reductase. *Science, 316*(5828), 1188-1191. doi:10.1126/science.1141179

Kato, T., Haruki, M., Imanaka, T., Morikawa, M., & Kanaya, S. (2001). Isolation and characterization of long-chain-alkane degrading Bacillus thermoleovorans from deep subterranean petroleum reservoirs. *J Biosci Bioeng, 91*(1), 64-70. doi:10.1263/jbb.91.64

Lee, Y-J., Park, M-K., Park, G-S., Lee, S-J., Lee, S. J., Kim, B. S., Shin, J-H., & Lee, D-W. (2017). Complete genome sequence of the thermophilic bacterium Geobacillus subterraneus KCTC 3922T as a potential denitrifier. *Journal of Biotechnology, 251*, 141-144. doi:https://doi.org/10.1016/j.jbiotec.2017.04.022

Levy-Booth, D. J., Navas, L. E., Fetherolf, M. M., Liu, L.-Y., Dalhuisen, T., Renneckar, S., Eltis, L. D., & Mohn, W. W. (2022). Discovery of lignin-transforming bacteria and enzymes in thermophilic environments using stable isotope probing. *The ISME Journal, 16*(8), 1944-1956. doi:10.1038/s41396-022-01241-8

Liang, J.-L., Jiangyang, J.-H., Nie, Y., & Wu, X.-L. (2016). Regulation of the Alkane Hydroxylase CYP153 Gene in a Gram-Positive Alkane-Degrading Bacterium, Dietzia sp. Strain DQ12-45-1b. *Applied and Environmental Microbiology, 82*(2), 608-619. doi:10.1128/aem.02811-15

Lin, J.-H., Zhang, K.-C., Tao, W.-Y., Wang, D., & Li, S. (2019). Geobacillus strains that have potential value in microbial enhanced oil recovery. *Applied Microbiology and Biotechnology, 103*(20), 8339-8350. doi:10.1007/s00253-019-10115-7

Liu, J., Zheng, Y., Lin, H., Wang, X., Li, M., Liu, Y., Yu, M., Zhao, M., Pedentchouk, N., Lea-Smith, D. J., Todd, J. D., Magill, C. R., Zhang, W-J., Zhou, S., Song, D., Zhong, H., Xin, Y., Yu, M., Tian J., & Zhang, X.-H. (2019). Proliferation of hydrocarbon-degrading microbes at the bottom of the Mariana Trench. *Microbiome, 7*(1). doi:10.1186/s40168-019-0652-3

Ma, B., Lu, C., Wang, Y., Yu, J., Zhao, K., Xue, R., Ren H., Lv, X. Pan, R., Zhang, J., Zhu, Y., Xu, J. (2023). A genomic catalogue of soil microbiomes boosts mining of biodiversity and genetic resources. *Nat Commun, 14*(1), 7318. doi:10.1038/s41467-023-43000-z

Mak, W. S., Wang, X., Arenas, R., Cui, Y., Bertolani, S., Deng, W. Q., Tagkopoulos, I., Wilson, D. K., & Siegel, J. B. (2020). Discovery, Design, and Structural Characterization of Alkane-Producing Enzymes across the Ferritin-like Superfamily. *Biochemistry, 59*(40), 3834-3843. doi:10.1021/acs.biochem.0c00665

Muhd Sakaff, M. K. L., Abdul Rahman, A. Y., Saito, J. A., Hou, S., & Alam, M. (2012). Complete Genome Sequence of the Thermophilic Bacterium Geobacillus thermoleovorans CCB_US3_UF5. *Journal of Bacteriology, 194*(5), 1239-1239. doi:10.1128/jb.06580-11

Okibe, N., Amada, K., Hirano, S., Haruki, M., Imanaka, T., Morikawa, M., & Kanaya, S. (1999). Gene cloning and characterization of aldehyde dehydrogenase from a petroleum-degrading bacterium, strain HD-1. *J Biosci Bioeng, 88*(1), 7-11. doi:10.1016/s1389-1723(99)80167-4

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology, 36*(10), 996-1004. doi:10.1038/nbt.4229

Que, Lawrence, & Dong, Y. (1996). Modeling the Oxygen Activation Chemistry of Methane Monooxygenase and Ribonucleotide Reductase. *Accounts of Chemical Research, 29*(4), 190-196. doi:10.1021/ar950146g

Rojo, F. (2009). Degradation of alkanes by bacteria. *Environmental Microbiology, 11*(10), 2477-2490. doi:10.1111/j.1462-2920.2009.01948.x

Roshick, C., Iliffe-Lee, E. R., & Mcclarty, G. (2000). Cloning and Characterization of Ribonucleotide Reductase fromChlamydia trachomatis. *Journal of Biological Chemistry, 275*(48), 38111-38119. doi:10.1074/jbc.m006367200

Ross, M. O., & Rosenzweig, A. C. (2017). A tale of two methane monooxygenases. *JBIC Journal of Biological Inorganic Chemistry, 22*(2-3), 307-319. doi:10.1007/s00775-016-1419-y

Serge Maurice Mbadinga and Li-Ying Wang and Lei Zhou and Jin-Feng Liu and Ji-Dong Gu and Bo-Zhong, M. (2011). Microbial communities involved in anaerobic degradation of alkanes. *International Biodeterioration & Biodegradation, 65*(1), 1-13. doi:https://doi.org/10.1016/j.ibiod.2010.11.009

Shafaat, H. S., Griese, J. J., Pantazis, D. A., Roos, K., Andersson, C. S., Popović-Bijelić, A., Gräslund A., Siegbahn, P. E. M., Neese, F., Lubitz, W., Högbom, M., & Cox, N. (2014). Electronic Structural Flexibility of Heterobimetallic Mn/Fe Cofactors: R2lox and R2c Proteins. *Journal of the American Chemical Society, 136*(38), 13399-13409. doi:10.1021/ja507435t

Sun, Y., Ning, Z., Yang, F., & Li, X. (2015). Characteristics of Newly Isolated Geobacillus sp. ZY-10 Degrading Hydrocarbons in Crude Oil. *Pol J Microbiol, 64*(3), 253-263.

Suzuki, H., Yoshida, K., & Ohshima, T. (2013). Polysaccharide-degrading thermophiles generated by heterologous gene expression in Geobacillus kaustophilus HTA426. *Appl Environ Microbiol, 79*(17), 5151-5158. doi:10.1128/aem.01506-13

Takami, H., Inoue, A., Fuji, F., & Horikoshi, K. (1997). Microbial flora in the deepest sea mud of the Mariana Trench. *FEMS Microbiology Letters, 152*(2), 279-285. doi:10.1111/j.1574-6968.1997.tb10440.x

Takami, H., Nishi, S., Lu, J., Shimamura, S., & Takaki, Y. (2004). Genomic characterization of thermophilic Geobacillus species isolated from the deepest sea mud of the Mariana Trench. *Extremophiles, 8*(5), 351-356. doi:10.1007/s00792-004-0394-3

Torrent, M., Musaev, D. G., Basch, H., & Morokuma, K. (2002). Computational studies of reaction mechanisms of methane monooxygenase and ribonucleotide reductase. *Journal of Computational Chemistry, 23*(1), 59-76. doi:10.1002/jcc.1157

Tully, B. J., Graham, E. D., & Heidelberg, J. F. (2018). The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. *Scientific Data, 5*(1), 170203. doi:10.1038/sdata.2017.203

Van Beilen, J. B., Li, Z., Duetz, W. A., Smits, T. H. M., & Witholt, B. (2003). Diversity of Alkane Hydroxylase Systems in the Environment. *Oil & Gas Science and Technology, 58*(4), 427-440. doi:10.2516/ogst:2003026

Voevodskaya, N., Lendzian, F., Ehrenberg, A., & Gräslund, A. (2007). High catalytic activity achieved with a mixed manganese–iron site in protein R2 of Chlamydia

ribonucleotide reductase. *FEBS Letters, 581*(18), 3351-3355. doi:10.1016/j.febslet.2007.06.023

Wang, W., & Shao, Z. (2013). Enzymes and genes involved in aerobic alkane degradation. *Front Microbiol, 4*, 116. doi:10.3389/fmicb.2013.00116

Wentzel, A., Ellingsen, T. E., Kotlar, H.-K., Zotchev, S. B., & Throne-Holst, M. (2007). Bacterial metabolism of long-chain n-alkanes. *Applied Microbiology and Biotechnology, 76*(6), 1209-1221. doi:10.1007/s00253-007-1119-1

Wushke, S., Levin, D. B., Cicek, N., & Sparling, R. (2015). Facultative Anaerobe Caldibacillus debilis GB1: Characterization and Use in a Designed Aerotolerant, Cellulose-Degrading Coculture with Clostridium thermocellum. *Applied and Environmental Microbiology, 81*(16), 5567-5573. doi:10.1128/aem.00735-15

Xu, X., Liu, W., Tian, S., Wang, W., Qi, Q., Jiang, P., Gao, X., Li, F., Li, H. & Yu, H. (2018). Petroleum Hydrocarbon-Degrading Bacteria for the Remediation of Oil Pollution Under Aerobic Conditions: A Perspective Analysis. *Frontiers in Microbiology, 9*. doi:10.3389/fmicb.2018.02885

Zhou, J-F., Gao, P-K., Dai, X-H., Cui, X-Y., Tian, H-M, Xie, J-J, Li, G-Q., & Ma, T. (2018). Heavy hydrocarbon degradation of crude oil by a novel thermophilic Geobacillus stearothermophilus strain A-2. *International Biodeterioration & Biodegradation, 126*, 224-230. doi:https://doi.org/10.1016/j.ibiod.2016.09.031