BACTERIAL DESULFONATION AND DEFLUORINATION OF

6:2 FLUOROTELOMER SULFONATE (6:2 FTS)

A Thesis

by

YING SHI

Submitted to the Office of Graduate and Professional Studies of Texas A&M University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Kung-Hui Chu	
Committee Members,	Libo Shan	
	Yina Liu	
Head of Department,	Robin Autenrieth	

December 2018

Major Subject: Civil Engineering

Copyright 2018 Ying Shi

ABSTRACT

6:2 Fluorotelomer sulfonate (6:2 FTS) is an emerging environmental contaminant. As a transformation product from the degradation of formulant in the aqueous film-forming foams (AFFF), 6:2 FTS has been frequently found with high concentrations in groundwater and soil, especially those adjacent to firefighting training areas. Due to its wide occurrence, toxicity, and bioaccumulation potential, 6:2 FTS has received a lot of attentions in the past decade. Conflicting biodegradability of 6:2 FTS under aerobic conditions were observed in activated sludge and river sediment. There was no evidence of biodegradation or biotransformation of 6:2 FTS occurring under anaerobic conditions.

Current knowledge on the factors determining biotransfomation and biodefluorination rate of 6:2 FTS is still unclear. To bridge this knowledge gap, this thesis characterized cultivable 6:2 FTS-degrading strains and elucidated the rate-limiting step affecting biotransfomation and biodefluorination of 6:2 FTS, with an emphasis on enzymes responsible for desulfonation. Two desulfonating enzyme systems, taurine dioxygenase (TauD) and two-component alkanesulfonate monooxygenase (SsuE/D), were examined for their ability to desulfonate 6:2 FTS.

A rhizosphere soil bacterial isolate, *Pseudomonas* strain SYC, can not only biotransform but also defluorinate 6:2 FTS. Two 6:2 FTOH-degrading strains, *Rhodococcus jostii* RHA1 and *Pseudomonas oleovorans*, also showed an ability to defluorinate 6:2 FTS. According to the degree of defluorination under different growth conditions, 6:2 FTS was readily defluorinated when it served as the sole sulfur source with appropriate carbon source being provided, such as ethanol, 1-butanol, and n-octane. There was no observable fluoride release when sulfate presents in the medium, likely due to the repression of the expression of desulfonating enzymes, suggesting that desulfonation is the first step in 6:2 FTS metabolism by 6:2 FTS-degrading strains.

Three desulfonating-associated enzymes, taurine dioxygenase (TauD), alkanesulfonate reductase (SsuE) and alkanesulfonate monooxygenase (SsuD), were successfully expressed and produced by *E. coli* competent cell BL21 (DE3). Free sulfite release was observed when using crude extract of enzymes to react with 6:2 FTS, indicating successful desulfonation by TauD and SsuE/D system.

The elucidations of rate-limiting step of 6:2 FTS defluorination, as well as enzymes responsible for 6:2 FTS desulfonation, provide fundamental knowledge for future studies on the molecular biology of 6:2 FTS metabolism in bacteria and the development of biological treatment strategies for enhanced 6:2 FTS removal.

ACKNOWLEDGMENTS

To my amazing advisor, Dr. Kung-Hui Chu: I wouldn't have gotten this far without your great mentorship and encouragement. Thank you for always being so patient, allowing me to grow, and giving me opportunities to challenge myself. Sincerely many Thanks!

I would like to express my appreciation to committee members, Dr. Libo Shan and Dr. Yina Liu, for guiding me with great advice and support throughout the course of this research. Thanks to you.

I must thank to my lab mates, Myung Hwangbo, Niao Yan, Chih-Hung Chen, Shih-Hung Yang, Yiru Shao, Fahad Asiri, Sherly Lie, Bilal Abada, Baixin Wang, for their help, advice, support, and friendship.

A special gratitude goes to my cheerleaders, Lang Shi, Casper Fu, Lan Huang and Norma Teetes: I miss our interesting chats. My caring and always enthusiastic friends. Thank you very much for all the help and encouraging words when I felt lost.

I am indescribably grateful to my father Biao Shi, my mother Weihong Li, and my boyfriend Eduardo Martinez. Thanks for your understanding, your unconditional love, never-ending encouragement, and for always being there for me. Your support has meant more to me than you could possibly realize.

And finally, last but by no means least, also to everyone that I met at Texas A&M: it was great sharing good memories with all of you during last three years.

Thanks for all your encouragement!

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Dr. Kung-Hui Chu of the Department of Civil Engineering, Dr. Libo Shan of the Department of Plant Pathology and Microbiology, and Dr. Yina Liu of the Geochemical and Environmental Research Group.

The data analyzed for Section 3.2.4 was provided by Dr. Yina Liu.

All other work conducted for the thesis was completed by the student independently.

Funding Sources

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

TABLE OF CONTENTS

	Page
ABSTRACT	ii
ACKNOWLEDGMENTS	iv
CONTRIBUTORS AND FUNDING SOURCES	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
1. INTRODUCTION AND OBJECTIVES	1
1.1. Introduction1.2. Goal, objectives, and hypotheses	1 4
2. LITERATURE REVIEW	7
 2.1. 6:2 Fluorotelomer sulfonate (6:2 FTS) 2.1.1. Physical and chemical properties of 6:2 FTS 2.1.2. Sources and fate of 6:2 FTS in the environments 2.2. 6:2 FTS biodegradation 	7 7 9
2.2. 6.2 FTS biodegradation in activated sludge and sediments 2.2.2. 6:2 FTS biodegradation by pure cultures	10
2.3. 6:2 FTOH-degrading strains 2.3.1. <i>Rhodococcus jostii</i> RHA1	14
2.3.2. <i>Pseudomonas oleovorans</i> 2.4. Microbial desulfonation	15
2.4.1. Bacterial responses to sulfur limitation2.4.2. Taurine dioxygenase2.4.3. Two-component alkanesulfonate monooxygenase system	16 18 19
3. MATERIALS AND METHODS	21
 3.1. Chemicals. 3.2. Isolation and characterization of 6:2 FTS-degrading bacteria. 	21
 3.2.1. Isolation of bacteria. 3.2.2. Defluorination and biotransfomation of 6:2 FTS by bacterial isolates 3.2.3. Sampling and sample preparation procedures 	

Page

3.2.4. HPLC/MS analysis	24
3.2.5. Taxonomic characterization of isolated bacteria	
3.3. Biodefluorination of 6:2 FTS by pure strains	
3.3.1. Bacterial strains and growth conditions	
3.3.2. 6:2 FTS defluorination with different growth substrates	
3.3.3. Fluoride measurement	
3.4. Assessment of desulfonation of 6:2 FTS by desulfonating enzymes	
3.4.1. Construction of expression plasmids	
3.4.2. Protein expression	30
3.4.3. Enzyme activity assay	
3.4.4. Desulfonation of 6:2 FTS by overexpressed enzymes	
J 1 J	
4. RESULTS AND DISCUSSION	
4.1. Identification and characteristics of a 6:2 FTS-degrading strain,	
Pseudomonas SYC	34
4.2. Factors affecting 6:2 FTS biodefluorination by pure strains	
4.2.1. Biodefluorination of 6:2 FTS by Rhodococcus jostii RHA1	
4.2.2. Biodefluorination of 6:2 FTS by Pseudomonas oleovorans	39
4.2.3. Biodefluorination of 6:2 FTS by Pseudomonas SYC	41
4.3. Desulfonation of 6:2 FTS by enzymes	42
4.3.1. Relationships of putative desulfonating bacteria	42
4.3.2. Cloning and overexpression of desulfonating enzymes in <i>E. coli</i>	44
4.3.3. Activities of enzymes	
4.3.4. Desulfonation of 6:2 FTS in crude cell extracts	46
5. SUMMARY, CONCLUSIONS AND FUTURE STUDIES	47
5.1. Summary and conclusions	
5.2. Future studies	49
REFERENCES	51
APPENDIX	63

LIST OF FIGURES

Figure 1.1. Structure of 6:2 fluorotelomer sulfonate (6:2 FTS)	1
Figure 2.1. Structure of perfluorooctane sulfonate (PFOS)	8
Figure 2.2. 6:2 FTS biotransformation pathway by <i>Gordonia</i> sp. NB4-1Y	13
Figure 2.3. Uptake and desulfonation of taurine and alkanesulfonates in <i>E. coli</i>	17
Figure 2.4. Mechanism for taurine desulfonation by taurine dioxygenase TauD	18
Figure 2.5. Mechanism of desulfonation by two-component alkanesulfonate monooxygenase system	20
Figure 3.1. Map of constructed recombinant plasmid pET11a-tauD	29
Figure 3.2. Map of constructed recombinant plasmids pET11a-ssuD and pET11a- ssuE.	30
Figure 4.1. Time course of biotransformation and defluorination of 6:2 FTS by the isolate strain <i>Pseudomonas</i> SYC	35
Figure 4.2. A phylogenetic tree showing relative relationship between strain SYC and related <i>Pseudomonas</i> species	36
Figure 4.3. <i>Rhodococcus jostii</i> RHA1 growth curve and time course of 6:2 FTS defluorination	39
Figure 4.4. <i>Pseudomonas oleovorans</i> growth curve and time course of 6:2 FTS defluorination	40
Figure 4.5. <i>Pseudomonas</i> SYC growth curve and time course of 6:2 FTS defluorination	41
Figure 4.6. Neighbor-joining tree illustrating the relationships of ssuD genes of putative desulfonating bacteria	42
Figure 4.7. Neighbor-joining tree illustrating the relationships of tauD genes of putative desulfonating bacteria	43
Figure 4.8. Overexpressed proteins of interest (SsuE, SsuD, and TauD)	45

1. INTRODUCTION AND OBJECTIVES

1.1. Introduction

6:2 Fluorotelomer Sulfonate (6:2 FTS, F(CF₂)₆CH₂CH₂SO₃H) (Fig. 1.1) is a polyfluorinated compound with six fluorinated carbons, two hydrocarbons and a sulfonate group. It is often utilized as mist suppressing agent in non-decorative hard chrome plating industry as an alternative to a banned chemical perfluorooctane sulfonate (PFOS).¹⁻⁵



Figure 1.1. Structure of 6:2 fluorotelomer sulfonate (6:2 FTS).

6:2 FTS has been found in surface water, groundwater, and soil, especially at firefighting training sites by US military, and locations where major fires have been extinguished through use of aqueous film-forming foams (AFFF).⁶ However, 6:2 FTS was not identified as a major component in AFFFs,⁷ suggesting that it may arise as intermediate transformation products from the degradation of precursor compounds,^{8,9} such as fluorotelomer thioether amido sulfonate (FtTAoS) that have been detected in AFFFs.¹⁰ In the past decade, a number of studies reported detections of 6:2 FTS ranging from ng/g to μg/g in soil^{6, 10, 11} as well as μg/L to mg/L in groundwater¹²⁻¹⁴ contaminated by AFFFs during historical fire-fighting training activities at airports or military bases,

the levels of which are comparable to those for perfluorooctanoate (PFOA) and PFOS.² The frequent detections of 6:2 FTS with high concentrations in the environment are quickly becoming a matter of concern.

6:2 FTS is classified as a Class III risk in a human risk study using the Threshold of Toxicological model with a threshold of 90 mg/person/day.² A recent report suggested that exposure to 6:2 FTS induced hepatotoxicity in the adult male mice, including liver weight increase, inflammation, and necrosis.³ After being exposed to 5 mg/kg/day of 6:2FTS for 28 days, 6:2 FTS was detected at high level in serum (18.52 µg/mL) as well as very high level in liver (194.44 µg/g), which is within the same order of magnitude as that of PFOA and PFOS, indicating its bioaccumulation potential and slow biodegradation in mice.³

Many advanced physical/chemical processes have shown effective removal of 6:2 FTS, such as nanofiltration,¹⁵ sonolysis,¹⁶ UV irradiation,⁴ electrochemical oxidation,^{17, 18} and advanced oxidation processes.^{1, 19} However, these techniques are costly and can potentially generate byproducts that are difficult to remove from water. On the other hand, biodegradation of 6:2 FTS by activated sludge, aerobic river sediments, and pure strains has been reported, suggesting that the biological approach could be an inexpensive and effective alternative for 6:2 FTS removal.

The biodegradability of 6:2 FTS under aerobic conditions were conflicting in activated sludge and river sediments. No evidence of biodegradation occurred under anaerobic conditions. The aerobic biotransformation process in activated sludge was extremely slow and incomplete, with 63.7% remaining at day 90.²⁰ However,

biotransformation of 6:2 FTS was relatively fast in the river sediment, with a half-life of approximately 5 days.²¹ It was speculated that the higher concentrations of monooxygenases, which may catalyze the enzymatic desulfonation process of 6:2 FTS, contributes to a faster rate of aerobic transformation.^{21, 22} In contrast, 6:2 FTS is persistent under anaerobic or anoxic conditions as suggested by the observation of no 6:2 FTS biotransformation in anaerobic river sediments over 100 days.²¹ These studies reported different extents of biodegradability of FTS under aerobic conditions. However, the conditions that caused the variation of biodegradation of FTS are unclear and have not been fully investigated.

Two pure cultures, *Gordonia* sp. NB4-1Y^{22, 23} and *Pseudomonas* sp. strain D2,²⁴ have shown the ability to use 6:2 FTS as their sole sulfur source under sulfur-limiting conditions. However, the enzyme responsible for the desulfonation of 6:2 FTS is not fully investigated.²⁵ Also, it is still unclear that if desulfonation of 6:2 FTS is a rate-limiting step of biodefluorination of 6:2 FTS. Given the high occurrence and concentration of 6:2 FTS in the environment and its toxicity,^{2, 20, 21} a better understanding on the mechanisms involved in factors and enzymes controlling the biotransformation of 6:2 FTS is thus warranted to the development of treatment for enhanced biodegradation of 6:2 FTS.

1.2. Goal, objectives, and hypotheses

The <u>goal</u> of this study is to investigate the rate-limiting step affecting biotransfomation and biodefluorination of 6:2 FTS, with an emphasis on enzymes responsible for desulfonation and defluorination. *The overall hypothesis is that the first step in 6:2 FTS metabolism in 6:2 FTS-degrading strains is the desulfonation of 6:2 FTS catalyzed by the expression of desulfonating-enzymes under sulfur-limiting growth conditions.* Below is the description of each sub-objective and associated subhypothesis.

Objective 1: Isolate and characterize 6:2 FTS-degrading bacteria from soils.

<u>Rationale and Hypothesis</u>: Many soil bacteria are capable of mineralizing sulfur from sulfate-esters and sulfonates, and these bacteria are commonly detected in the rhizosphere zone,²⁶ where biodegradation of organic pollutants including sulfate esters and sulfonates mainly occurs in soil.²⁷ Sodium dodecyl sulfate (SDS, CH₃(CH₂)₁₁SO₄⁻) is a main component in the detergent of car wash.²⁸ The desulfonating-associated microbial populations are enriched in the rhizosphere of soils that are contaminated with detergent at car wash facilities.^{28, 29} Accordingly, *I hypothesize that rhizosphere bacteria at car wash sites are capable of desulfonating 6:2 FTS*.

Task 1a: Enrich and isolate desulfonating bacteria from soil rhizosphere.

Task 1b: Examine capability of metabolizing 6:2 FTS by the isolated strains.

Task 1c: Examine capability of defluorination of 6:2 FTS by the isolates.

Objective 2: Examine the biodegradation potential of 6:2 FTS by 6:2 FTOHdegrading strains.

<u>Rationale and Hypothesis</u>: The structure of 6:2 FTS is analogous to that of 6:2 fluorotelomer alcohol ($F(CF_2)_6CH_2CH_2OH$, 6:2 FTOH), which was degraded under aerobic conditions in numerous environmental matrices^{30, 31} as well as by several pure cultures.^{32, 33} The primary biotransformation of 6:2 FTS was proposed to form 6:2 FTUA ($F(CF_2)_5CF=CHCOOH$), which was subsequently transformed following the pathways similar to that of 6:2 FTOH.²⁰ Accordingly, it is possible that the 6:2 FTOH degraders may also degrade 6:2 FTS. To test this hypothesis, two 6:2 FTOH-degrading strains, *Rhodococcus jostii* RHA1 and *Pseudomonas oleovorans*,³² were used.

Task 2a: Examine the biodefluorination of 6:2 FTS by *Rhodococcus jostii* RHA1 under different oxygenase-expression conditions.

Task 2b: Examine the biodefluorination of 6:2 FTS by *Pseudomonas oleovorans* under different oxygenase-expression conditions.

Objective 3: Examine enzymes responsible for the desulfonation of 6:2 FTS.

<u>Rationale and Hypothesis</u>: Previous studies have shown that the 6:2 FTS is susceptible to biodegradation under aerobic, sulfur-limiting conditions. It was speculated that 6:2 FTS must be desulfonated first for further biotransformation to occur.^{20, 21} Many studies have reported that aerobic bacteria liberate sulfite from sulfonates by using oxygenases, such as α -ketoglutarate-dependent taurine dioxygenase (TauD) and an FMNH₂-dependent

alkanesulfonate monooxygenase system (SsuE/D).³⁴⁻³⁶ The structure of two hydrocarbons with a sulfonate moiety of 6:2 FTS is similar to those of many aliphatic sulfonates. *It is hypothesized that taurine dioxygenase (TauD) and two-component alkanesulfonate monooxygenase system (SsuE/D) may be responsible for 6:2 FTS desulfonation.*^{22, 37}

Task 3a: Conduct cloning by constructing expression plasmids of flavin reductase gene (ssuE), alkanesulfonate monooxygenase gene (ssuD), and taurine dioxygenase gene (tauD) and overexpress proteins.

Task 3b: Examine the capabilities of alkanesulfonate monooxygenase system (SsuE/D) and taurine dioxygenase (TauD) on 6:2 FTS desulfonation.

2. LITERATURE REVIEW

2.1. 6:2 Fluorotelomer sulfonate (6:2 FTS)

2.1.1. Physical and chemical properties of 6:2 FTS

Due to the limited physical/chemical property data for 6:2 FTS, most properties for the free acid forms of 6:2 fluorotelomer sulfonate are estimated (Table. 2.1) using property estimation software.² 6:2 FTS is white to off-white crystalline powder without sensible odor. The chemical formula for 6:2 fluorotelomer sulfonic acid is $C_8H_5F_{13}O_3S$ and its molecular weight is 428.17 g/mol.

Predicted property	Value
CAS Number	27619-97-2
Chemical Formula	$C_8H_5F_{13}O_3S$
Molecular Weight (g/mol)	428.17
рКа	$1.31 \pm 0.50, 0.36$
log Vapor pressure (Pa)	-0.96
log water solubility (mol/L, 25 °C)	-2.51
log K _{ow} (octanol-water coefficient, 25 °C)	4.44; 2.455
log K _{aw} (air-water coefficient)	-4.85
log K _{oa} (octanol-air coefficient)	9.28
K _d (solid-water partition coefficient, L/kg)	3.1 - 12
Koc (soil-organic carbon coefficient, pH 7, 25	1.0
BCF (bioconcentration factor)	1.0
LC ₅₀ mg/kg (oral in Sprague Dawley rats)	1,871

Table 2.1. Predicted properties of the free acid forms of 6:2 Fluorotelomer Sulfonate (6:2 FTS).²

6:2 FTS is thermally stable, which has a melting point over 300 °C. With estimated pKa of 0.36–1.31, 6:2 FTS presents primarily in the anionic form at environmentally relevant pH. 6:2 FTS is relatively non-volatile and non-soluble in water, but readily soluble in organic solvents such as ethanol.

The chemical structure of 6:2 FTS is similar to that of perfluorooctane sulfonate (PFOS) (Fig. 2.1). In the hard metal plating industry, 6:2 fluorotelomer sulfonate is usually applied as the alternative to perfluorooctane sulfonate (PFOS), which has been listed in the Stockholm Convention for global ban in the early 2000s due to its potential hazards and persistent nature.^{1, 4, 9, 38}



Figure 2.1. Structure of perfluorooctane sulfonate (PFOS).

It should be noted that 6:2 FTS has been referred to in some literature as "H-PFOS",²⁴ "THPFOS",³⁹ "tetrahydro PFOS",⁶ but 6:2 FTS is both physically and chemically different from PFOS.²⁰ 6:2 FTS has relatively lower acidity and surface activity compared to perfluorinated acids.⁴⁰ Due to its slightly higher surface tension, 6:2 FTS can only be partly applied in decorative plating.³⁸ The sorption of several fluorotelomers to six soil samples with varying soil properties and characteristics was evaluated.⁴¹ Among anionic fluorotelomer sulfonates (FTSs), zwitterionic fluorotelomer sulfonamido betaines (FtSaBs), and cationic 6:2 fluorotelomer sulfonamido amine (FtSaAm), 6:2 FTS was the weakest FTSs sorbed to soil.⁴¹ The low K_d values for the 6:2 FTS suggested that it is highly mobile in groundwater,⁴¹ which is in accord with the high concentrations of 6:2 FTS in the analysis of AFFF-impacted groundwater (up to 14.6 mg/L).¹²

2.1.2. Sources and fate of 6:2 FTS in the environments

Widespread detections of 6:2 FTS in a variety of environmental media were observed. 6:2 FTS has been found in groundwater (up to 14.6 mg/L at AFFF-impacted area and 8.54 ng/L in non-industrial area),^{12,38} rivers (1.6 ng/L in the Hai River, China),⁴² wastewater effluents (median concentration of 3.4 ng/L from eight WWTP in San Francisco Bay),⁴³ and soil (612-2101 ng/g at Flesland airport, Norway).¹¹ They have also been detected in municipal sewage sludge (13.9 ng/g, China),⁴⁴ sediments (7 ng/g at Flesland airport, Norway),¹¹ WWTP effluents (37.9 ng/L, Germany),⁴⁵ landfill leachate (582 ng/L),⁴⁶ and human serum(0.047 μ g/L).⁴⁷ The surface water, sediment, groundwater, and soil with high reported levels of 6:2 FTS are adjacent to firefighting training areas associated with repeated usage of AFFF use and systems impacted by fluorinated chemical manufacturing wastes.^{12, 48}

The primary origin and environmental fate of 6:2 FTS could be degradation or transformation of more complex fluorotelomer-based precursors containing the $C_nF_{2n+1}CH_2CH_2S-R$ or $C_nF_{2n+1}CH_2CH_2SO_2-R$ moiety (where R is a hydrophilic functional group that provides surfactant properties)^{12, 44} in the firefighting foams.

Recent studies reported that several formulations of AFFFs, branded as Angus-fire Tridol, Ansulite, and Angus-fire Tridol, were evaluated. The concentration of 6:2FTS was found to have significantly increased after oxidization of AFFF formulations, although it was not obvious before oxidization.⁹ For example, Fluorotelomer thioether amido sulfonate (FtTAoS) is present in several widely used AFFF formulations,¹⁰ which may undergo biodegradation or biotransformation to 6:2 FTS in soil, groundwater or other environments where AFFFs were released.

2.2. 6:2 FTS biodegradation

2.2.1. 6:2 FTS biodegradation in activated sludge and sediments

Aerobic degradation of 6:2 FTS has been observed in activated sludge (sampled from three wastewater treatment plants (WWTPs) in Pennsylvania, Maryland, and Delaware) and river sediments (sampled from Brandywine river creek, Pennsylvania).

6:2 FTS was readily biotransformed in aerobic river sediment,²¹ slowly and incompletely transformed in activated sludge,²⁰ but was not biotransformed in anaerobic river sediment.²¹ The half-life of the biotransformation process in activated sludge and aerobic sediment are approximately 2 years and 5 days,^{20, 21} respectively. The significant difference of aerobic transformation rate might be due to the great disparities in the concentrations of monooxygenase, which may be responsible of catalyzing the enzymatic desulfonation process of 6:2 FTS.²¹ Primary biotransformation of 6:2 FTS after 90 days in activated sludge bypassed 6:2 fluorotelomer alcohol (6:2 FTOH, F(CF₂)₆CH₂CH₂OH) to form 6:2 FTUA directly, which was subsequently transformed following the pathways similar to that of 6:2 FTOH to form PFPeA (F(CF₂)₄COOH) and PFHxA (F(CF₂)₅COOH) eventually.²⁰ The biotransformation of 6:2 FTS after 90 days in aerobic river sediment proceeded with the formation of 5:3 fluorotelomer carboxylic acid (5:3 FTCA) (F(CF₂)₅CH₂CH₂COOH, 16 mol%), PFPeA (21 mol%), and PFHxA (20 mol%).²¹ Compared to the parallel set of 6:2 FTOH biotransformation in the same aerobic sediment, 6:2 FTS primary biotransformation and subsequent formation of stable transformation products was as fast as that of 6:2 FTOH.²¹ In contrast, 6:2 FTOH was readily biotransformed to 6:2 FTCA (60 mol%) and 5:3 FTCA (12 mol%) whereas 6:2 FTS was persistent in anaerobic sediments over 100 days. This distinctive difference, compared with 6:2 FTOH, is attributable to the sulfonate moiety of 6:2 FTS. The biotransformation performance of 6:2 FTS and 6:2 FTOH under three conditions and the analogous structure indicated that 6:2 FTS desulfonation is the initial step required for 6:2 FTS to be further biotransformed to downstream stable transformation products.^{20, 21}

2.2.2. 6:2 FTS biodegradation by pure cultures

Relatively few studies investigated the biodegradation of 6:2 FTS in aqueous phase with pure strains. To date, two pure cultures, *Gordonia* sp. NB4-1Y^{22, 23} and *Pseudomonas* sp. strain D2,²⁴ were reported capable of biodegrading and biodefluorinating 6:2 FTS under aerobic conditions, in which the 6:2 FTS was the sole sulfur source.

Pseudomonas sp. strain D2 can defluorinate 6:2 FTS into six volatile transformation products. All of the volatile products contained carbon, hydrogen, oxygen

and fluorine, but not sulfur. This suggests a linkage between sulfur assimilation and defluorination. The incubations closed to atmosphere and open to atmosphere were compared. Nearly 1.42 mol of fluoride release was detected per mole of 6:2 FTS added in closed incubations. However, this ratio was only 0.09 mol F⁻/mol 6:2 FTS in the open incubations, indicating further defluorination of volatile intermediates in the closed systems.²⁴

A strain isolated from vermicompost, *Gordonia* sp. NB4-1Y, can metabolize 6:2 FTS when 6:2 FTS was supplied as sole sulfur source.²² The strain degraded 6:2 FTS (60 μ M) into ten major breakdown products over 7 days under aerobic condition.²³ Greater than 50% of transformed products were found in volatile phase. About 20% of the total 6:2 FTS spiked was transformed to water-soluble metabolites. Based on the detected degradation products, two pathways (major 5:2 ketone and minor 5:3 FTCA) for 6:2 FTS biodegradation by NB4-1Y were proposed (Fig. 2.2).²³

Both strains were capable of degrading 6:2 FTS only under sulfur-limiting conditions. The presence of inorganic sulfur, such as sulfate and sulfite, might repress the expression of enzymes for sulfur removal from organic molecules.²²

12



Figure 2.2. 6:2 FTS biotransformation pathway by *Gordonia* sp. NB4-1Y.²³ Reprinted from [23].

2.3. 6:2 FTOH-degrading strains

The chemical structure of 6:2 FTS is very similar to 6:2 fluorotelomer alcohol (6:2 FTOH, F(CF₂)₆CH₂CH₂OH), where 6:2 FTS contains a sulfonate end (-SO₃⁻) and 6:2 FTOH contains a hydroxyl end (-OH). It has been proposed that 6:2 FTOH would be formed and easily biotransformed by 6:2 FTOH-degrading strains if the sulfonate group of 6:2 FTS is removed and substituted with a hydroxyl group.²¹ *Pseudomonas oleovorans* and *Rhodococcus jostii* RHA1 are two well-studied bacteria known for their ability to biotransform FTOHs or polychlorinated compounds,^{32, 49} and aliphatic compounds. These two strains can grow on a wide range of carbon sources and express different oxygenases in responding to their growth substrates. Thus, these two strains are selected for their ability to degrade 6:2 FTS in this study.

2.3.1. Rhodococcus jostii RHA1

Rhodococcus jostii strain RHA1 (referred as RHA1 hereafter) is a gram-positive strain, which was isolated from γ-hexachlorocyclohexane-contaminated soil in Japan in 1995.⁵⁰ RHA1 is capable of catabolizing a wide range of compounds including polychlorinated biphenyls (PCB), trichloroethylene (TCE), 1,4-dioxane and triclosan.⁴⁹⁻⁵¹ The complete genome sequence of RHA1 reveals that RHA1's protein-encoding genes are outstandingly rich in oxygenases.⁵² Many of the oxygenases expressed by RHA1, such as propane monooxygenase (PMO) and soluble butane monooxygenase (sBMO), played important roles in the numerous degradation pathways of polychlorinated or polyfluorinated compounds.^{49, 51} Based on the BLAST result, RHA1 genome also

contains the genes encoding taurine dioxygenase (TauD), flavin reductase (SsuE), and alkanesulfonate monooxygenase (SsuD).⁵³ Several *Rhodococcus* strains were reported to cleave C-S bond of sulfonates to supply sulfur for bacterial growth, including RHA1.⁵³

2.3.2. Pseudomonas oleovorans

Pseudomonas oleovorans (referred as *P. ole* hereafter) is a gram-negative strain isolated from soil in 1963.⁵⁴ *P. ole* is able to use various C_6 - C_{12} linear alkanes or alkenes as the sole carbon and energy sources on account of possessing an OCT plasmid, which encodes a complex of monooxygenases.⁵⁵ The OCT plasmid consists the *alkBAC* operon and the *alkR* regulatory region. The *alkBAC* operon encodes alkane hydroxylase (a threecomponent monooxygenase adding one oxygen atom to a substrate), alkanol dehydrogenases and aldehyde dehydrogenases.^{56, 57} These enzymes allow *P. ole* to oxidize alkanes to alkanols, alkanals and alkanoic acids.

Pseudomonas oleovorans was reported to successfully covert 6:2 FTUCA, intermediate biostransformed product of 6:2 FTOH, to 5:2 ketone through a major degradation pathway and to 5:3 FTUA through a minor degradation pathway.³²

2.4. Microbial desulfonation

Sulfur is essential for all living organisms, since it is required in the composition of amino acids cysteine and methionine. It is also an important constituent of vitamins (biotin and thiamine) and various enzyme cofactors such as coenzyme A, thioredoxin, and glutathione.³⁴ Sulfur makes up 0.5-1% of the cell dry weight in bacteria.³⁴ Bacteria

can derive sulfur from different sulfur sources, including inorganic sulfate, sulfite, cysteine, thiosulfate, and thiocyanate.^{34, 58-60} However, the preferred sulfur source is species dependent.

Sulfonates, with general formula R-C-SO3⁻, are chemically stable compounds. They are common xenobiotics in the environment.^{34, 61} The naturally occurring sulfonates were considered to be relatively few in number, though often in significant contributions to the global biogeochemical sulfur cycle.⁶¹ When the preferred forms of sulfur element are not available, many bacterial species can utilize organosulfur compounds including sulfate esters and sulfonates as sulfur supply.^{58, 60}

2.4.1. Bacterial responses to sulfur limitation

A number of "sulfate starvation-induced proteins" (SSI proteins), involved in the uptake and desulfurization of sulfonates and sulfate esters, has been found in several Gram-positive and Gram-negative species, such as *Escherichia coli*, *Pseudomonas aeruginosa*, and *Rhodococcus* sp. SY1.^{60, 62, 63} However, many of these enzyme systems involved in organosulfur assimilation have been reported to be subject to negative regulation by the presence of sulfate and cysteine.⁵⁹ Bacteria synthesize SSI proteins for adaptation to sulfur-limiting conditions only when cells are starved of preferred sulfur sources. This sulfate starvation response is essential for bacteria to survive in the natural environment.⁵⁸

The mechanism and genes involved in organosulfur metabolism have been elucidated in several strains including *Escherichia coli, Bacillus subtilis, Pseudomonas* *aeruginosa*, and *Pseudomonas putida*.^{34, 58} Many studies have shown that aerobic bacteria liberate sulfite from sulfonates by using oxygenases. Two main enzymatic desulfonation routes have been characterized, an α -ketoglutarate-dependent dioxygenase pathway that is active with taurine and alkanesulfonates, and an FMNH₂-dependent monooxygenase system that catalyzes the desulfonation of a range of aliphatic sulfonates (Fig. 2.3).³⁴⁻³⁶ Both systems employ ABC-type transporters involved in the cellular uptake of taurine or alkanesulfonates. These two enzymatic pathways work in parallel to ensure that the bacterial cell has sufficient sulfur for biosynthetic processes.⁶⁴



Figure 2.3. Uptake and desulfonation of taurine and alkanesulfonates in *E. coli*.³⁵ Reprinted from [35].

2.4.2. Taurine dioxygenase

The taurine dioxygenase (TauD) is a nonheme iron(II) and α -ketoglutarate (α KG) dependent dioxygenase, which was identified during screening for *Escherichia coli* proteins.⁶⁵ The synthesis of TauD was expressed during growth with alkanesulfonates under sulfur-limiting condition, but repressed in the presence of sulfate.^{34, 59} TauD converts taurine to sulfite and aminoacetaldehyde in the presence of oxygen, α KG, and Fe(II) as illustrated in Fig. 2.4.^{66, 67}



Figure 2.4. Mechanism for taurine desulfonation by taurine dioxygenase TauD.⁶⁷ Reprinted from [67].

The reaction consumes equimolar amounts of oxygen and α KG. Ferrous iron is indispensable for activity and ascorbate leads to a 50% enhance of the activity.²⁵ Taurine dioxygenases transfer both oxygen atoms to organic acceptor molecules.³⁴ One oxygen atom hydroxylates the reaction substrate (taurine or alkanesulfonates) to form

intermediates, which is unstable and spontaneously decomposes to yield final products,^{66, 67} the other is transferred to the cosubstrate α KG, which is consequently converted to succinate and carbon dioxide.³⁴ Taurine is the preferred substrate by TauD, but pentanesulfonate, MOPS (3-(*N*-morpholino)propanesulfonic acid), butanesulfonate, and 1,3-dioxo-2-isoindolineethanesulfonic acid also can be desulfonated by TauD at significant rates.²⁵

2.4.3. Two-component alkanesulfonate monooxygenase system

The two-component alkanesulfonate monooxygenase system enables bacteria to utilize diverse alkanesulfonates as alternative sulfur sources.⁶⁸ This system comprises of a nicotinamide adenine dinucleotide phosphate (NADPH)-dependent flavin mononucleotide (FMN) reductase, SsuE, and an FMNH₂-dependent alkanesulfonate monooxygenase, SsuD.^{34, 69} Unlike a traditional flavin monooxygenase, SsuD is only involved in the oxidative half-reaction, instead of both oxidative and reductive half-reactions occurring on the same enzyme.⁶⁸ In the overall reaction, SsuE catalyzes the reduction of FMN directly by NADPH to form reduced flavin FMNH₂ which is subsequently transferred to SsuD (Fig. 2.5).³⁶ The FMNH₂ binds to SsuD enzyme and activates molecular oxygen to form a C4a-(hydro)peroxyflavin intermediate (FMNO⁻). The C4a-(hydro)peroxyflavin performs a nucleophilic attack on the sulfur atom of the alkanesulfonate substrate followed by a Baeyer-Villiger rearrangement leading to the cleavage of the carbon-sulfur bond, releasing corresponding aldehyde and sulfite.^{68, 70} The sulfite product is then assimilated for the survival and growth of bacteria.⁷¹



Figure 2.5. Mechanism of desulfonation by two-component alkanesulfonate monooxygenase system.³⁶ Reprinted from [36].

It appears that the SsuE is not an essential part in the catalyzation of desulfonation; its role is limited to providing FMNH₂ for the SsuD.³⁴ SsuD or SsuD-like enzymes haven been widely found in many bacterial strains such as *Bacillus subtilis*, *Pseudomonas* strains, and *Corynebacterium glutamicum*.^{37, 62, 72} Similar to TauD, the expression of the ssu gene is repressed in the presence of inorganic sulfate, sulfite, sulfide, or cysteine.³⁴ Numerous sulfonates have been successfully desulfurized by two-component alkanesulfonate monooxygenase system including propanesulfonate, octanesulfonate, *N*-Phenyltaurine, Phenyl-1-butanesulfonate, hexanesulfonate, HEPES (4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid), MOPS (3-(N-morpholino)-propanesulfonic acid), PIPES (piperazine-*N*,*N*-bis(2-ethanesulfonic acid)), 2-(4-Pyridyl)ethanesulfonate, 1,3-Dioxo-2-isoindolineethanesulfonate.⁷³

3. MATERIALS AND METHODS

3.1. Chemicals

1H, 1H, 2H, 2H-Perfluorooctanesulfonic acid (6:2 FTS, CAS# 27619-97-2, 98% pure) was purchased from Synquest Laboratories (Alachua, FL). Riboflavin 5'monophosphate sodium salt (FMN, 93% pure) and 1-butanol (99.4% pure) were purchased from Fisher Scientific (Fair Lawn, NJ). Glycerol, n-octane (97% pure), and iron (II) chloride tetrahydrate were from Acros Organics (Geel, Belgium). Ellman's reagent (5,5'-Dithiobis (2-nitrobenzoic acid)), taurine, and acetonitrile (99.8% pure, HPLC grade) were purchased from Alfa Aesar (Ward Hill, MA). Ethanesulfonic acid sodium salt monohydrate (97% pure), NADPH tetrasodium salt, and α -ketoglutaric acid disodium salt dehydrate (98% pure) were purchased from Sigma-Aldrich (St. Louis, MO). Sodium ascorbate, isopropyl- β -D-1-thiogalactopyranoside (IPTG) and FastDNA Kit were purchased from MP Biomedicals (Santa Ana, CA). 6:2 FTS was prepared in absolute ethanol as a stock solution with a final concentration 5 g/L.

3.2. Isolation and characterization of 6:2 FTS-degrading bacteria

3.2.1. Isolation of bacteria

Bacteria capable of utilizing 6:2 FTS as a sole sulfur source were isolated from rhizosphere soils contaminated with detergent from car wash outlets in Bryan, Texas. Five grams of soil samples were inoculated into 250 mL Erlenmeyer flasks containing 50 mL sulfur-free basal salt media with the supplement of 1 g/L glucose and 0.3 g/L

ethanesulfonic acid sodium salt as sole sulfur source. The sulfur-free basal salts medium was prepared with the following ingredients (grams per liter): K_2HPO_4 , 3.5; KH_2PO_4 , 1.5; NH_4Cl , 0.5; $MgCl_2 \cdot 6H_2O$, 0.15; and NaCl, 0.62. The medium also contained the following trace elements (grams per liter): $FeCl_3 \cdot 6H_2O$, 0.24; $CoCl_2 \cdot 6H_2O$, 0.04; $CuCl_2 \cdot H_2O$, 0.041; $MnCl_2 \cdot 4H_2O$, 0.03; $ZnCl_2$, 0.147; and $Na_2MoO_4 \cdot 2H_2O$, 0.03.

The cultures were incubated at 30°C with constant shaking in at 150 rpm. After 7 days, one milliliter of enrichment was subcultured to 50 mL fresh sulfur-free basal medium. Following four times of weekly subcultures, the sulfur source was switched to 100 ppm 6:2 FTS (final concentration) and carbon source switched to 2% ethanol (v/v). Ethanol was used because 6:2 FTS stock solution was prepared in ethanol. After three weekly subcultures, 1 mL liquid culture was taken from the final enrichment flask, serially diluted, and transferred to sulfur-free basal medium agar plate containing 100 ppm 6:2 FTS and 2% ethanol. The strains were subsequently isolated by routinely streaking on sulfur-free basal salt medium agar plate to purity. The pure cultures were preliminarily screened for their ability to grow on 6:2 FTS as a sole sulfur source.

3.2.2. Defluorination and biotransfomation of 6:2 FTS by bacterial isolates

Three 6:2 FTS-utilizing isolates were selected for the defluorination screening. A single colony of each isolate was inoculated into 120 mL serum bottle containing 15 mL sulfur-free basal medium with the addition of 40 μ M 6:2 FTS as sole sulfur source and 0.3% ethanol. The 120-mL glass bottles were sealed with butyl rubber septa and crimped with aluminum caps. Cell-free controls were used to ensure study integrity. All samples

were in duplicate. The bottles were incubated on a shaker at 150 rpm in a 30°C dark room. The bottles were sacrificed over different time points: 0, 1, 3, 5 and 10 d for 6:2 FTS biotransformation analysis and fluoride release analysis.

3.2.3. Sampling and sample preparation procedures

At sampling time, prior to SPE extraction, an air pump (2.6 L min⁻¹ rated flow rate) was used to purge the headspace for 3 min to capture volatile transformation products into the C18 cartridges. The C18 cartridges were subsequently removed and eluted with 5 mL acetonitrile into 10 mL glass vials. The spent septa from each bottle was placed into a new glass bottle and extracted with 5 mL acetonitrile for 2-3 days at 50 °C.^{20, 21, 32} SPE cartridges (Oasis Wax[®] 200 mg/6 mL; Waters Corporation, Milford, MA) were used to extract samples according to previously described SPE clean up procedures with some modifications.^{15, 74, 75} The cartridges were conditioned using 5 mL of methanol with 0.1% (v/v) NH₄OH, 5 mL of methanol and 5 mL of water and dried completely under vacuum. The non-filtered water samples were passed through cartridges at a flow rate of 3-4 mL/min. The analytes were eluted with 1 mL of methanol, 4 mL of methanol with 0.1% NH₄OH and then 2 mL of 70:30 dichloromethane: isopropanol with 0.1% NH₄OH. The extract from the spent septa was pooled together with the extract from the SPE eluent. The combined extract was evaporated to dryness under a gentle stream of nitrogen and the residue reconstituted in 100 µL of 50:50 water/methanol and preserved in a freezer (-50°C) until being analyzed by HPLC/MS.

3.2.4. HPLC/MS analysis

To quantify 6:2 FTS in liquid samples, HPLC/MS analysis was performed using a High-Performance Liquid Chromatography (HPLC, Agilent 1290 Infinity II) / Triple Quadrupole Mass Spectrometer (QqQ-MS, Agilent 6470) equipped with a Jet Stream electrospray ionization (ESI) source. Briefly, 10μ L of samples were injected and then separated by an Agilent ZORBAX Eclipse Plus C-18 narrow bore (2.1mm×100mm, 1.8µm) HPLC column maintained at 50°C. HPLC flow rate was 0.4 mL/min. Chromatographic separation was achieved on a Solvent A (5mM ammonium acetate in water), and Solvent B (95% MeOH and 5% water with 5mM ammonium acetate). The separation gradient method used was: 0-0.5 min (holding at 10% B), 0.6-2 min (10% B to 30% B), 2.1-14 min (30% B to 95% B), 14.1-14.5 min (95 % B to 100 % B), 14.6 to 16.5 min (holding at 100% B), and then stabilize column at 10% B for 6 min before the next injection. The MS parameters were optimized for 6:2 FTS under direct infusion at 0.4 mL/min to identify the MRM transitions (precursor/product fragment ion pair). Sample acquisition and analysis were performed with MassHunter B.08.02 (Agilent).

3.2.5. Taxonomic characterization of isolated bacteria

Phylogenetic identification of the isolate was performed by 16S rRNA gene analysis. Cells were grown in 25 mL nutrient broth at 30°C overnight and harvested by centrifugation at 10,000 g for 10 min. Genomic DNA was extracted using the FastDNA Kit. The 16S rRNA gene was amplified by PCR with universal primers 27F (5'-AGA GTT TGA TCM TGG CTC AG-3') and 1492R (5'-GGT TAC CTT GTT ACG ACT T- 3'). The PCR program consisted of initial denaturation 95°C for 10 min, followed by 30 cycles each of 95°C for 1 min, 53°C for 1 min, 72°C for 40 s; 72°C for 10 min; and 4°C for 10 min. The PCR product was run on agarose gel and purified by using a QIAquick PCR Purification Kit (Qiagen). 16S rRNA gene sequencing of the isolate was carried out by Eton Bioscience, Inc. Sequences were compared to known 16S rRNA gene sequences in the NCBI database using the BLAST program and aligned using the ClustalW program. A phylogenetic tree (neighbor-joining method) was constructed using the MEGA7 software with 1000 bootstraps.

3.3. Biodefluorination of 6:2 FTS by pure strains

In this study, the effects of different growth substrates (1-butanol, ethanol, taurine, n-octane) on 6:2 FTS degradation potential of strains *Rhodococcus jostii* RHA1, *Pseudomonas oleovorans* as well as the 6:2 FTS-degrading isolate, named as *Pseudomonas* SYC, were investigated. The biodegradation potential of each strain grown under different growth conditions was evaluated in terms of the degree of defluorination.

3.3.1. Bacterial strains and growth conditions

Rhodococcus jostii RHA1 was grown in AMS medium⁵¹ with ethanol (0.3%, v/v) to an optical density OD_{600} of 0.5-0.7 before harvested for experimental use. The cells were centrifuged at 10,000 g for 20 min, and then the pellet was washed and resuspended in sulfur-free basal salt medium for experimental use. The cell suspensions

of *Pseudomonas oleovorans* and *Pseudomonas* SYC were prepared similarly to *Rhodococcus jostii* RHA1, except that *Pseudomonas oleovorans* was pregrown in AMS medium with 5% n-octane.

3.3.2. 6:2 FTS defluorination with different growth substrates

Experiments for determining 6:2 FTS defluorination degree were performed in a series of 120-mL glass serum bottles containing cell suspension (15 mL), 6:2 FTS (40 μ M final concentration) and one type of carbon source (referred as DF batch hereafter). To remove the solvent ethanol from 6:2 FTS stock solution, 51.5 µL of 6:2 FTS stock solution (5 g/L) was spiked into each serum bottle and put in a biosafety cabinet overnight to evaporate ethanol. 15 mL sulfur-free basal salt medium were added into each serum bottle supplement with one type of carbon source to bring the initial concentration of carbon to 110-120 mM. 200 µL of pre-grown cultures were spiked to each serum bottle at time 0. Butyl rubber septum and aluminum cap were used to seal each bottle. The effect of carbon source on 6:2 FTS biotransformation by Rhodococcus jostii RHA1 was tested with 1-butanol, ethanol, taurine. Pseudomonas oleovorans was tested with n-octane, ethanol, taurine. Pseudomonas SYC was tested with 1-butanol, ethanol, and taurine. Sulfur-contained controls were also used to ensure study integrity. The sulfur-contained basal salt medium recipe was prepared by replacing NH₄Cl, MgCl₂, and NaCl to (NH₄)₂SO₄, MgSO₄, and Na₂SO₄ with the same molar concentration of cation. The sulfur-contained controls were set up using sulfur-contained basal medium with ethanol as carbon source for RHA1, SYC, and P. ole. A parallel set of bottles was

used to monitor the optical density (OD_{600}) change with time (referred as OD batch hereafter). All samples and controls were in duplicate. The serum bottles were incubated on a shaker at 150 rpm in a 30°C dark room. The DF batch samples were sacrificed over time points according to the parallel OD batch measurement result: early exponential phase, the middle of exponential phase, and stationary phase.

3.3.3. Fluoride measurement

As 6:2 FTS is a polyfluorinated compound, the amount of fluoride released from 6:2 FTS biodegradation can be used to assess the degree of defluorination. The fluoride concentration of DF batch at each sampling time was measured after extracting volatile metabolites. The aqueous samples were first centrifuged at 13,000 rpm for 5 min and filtered through a 0.45 µm-pore size filter to remove the cell debris before measurement. The fluoride ion concentrations were determined using an Orion 96-09BNWP ion-selective electrode (Thermo Scientific, Beverly, MA) following the manufacturer's protocol. Fluoride standard solutions, ranging from 25-2500 ppb, were prepared by diluting a certified fluoride standard solution (100 mg/L, Thermo Scientific, Beverly, MA) in the corresponding TISAB II solution. Readings were recorded as milivolts and plotted against known fluoride concentrations to develop a linear standard calibration curve. The detection limit of fluoride concentration is 25 ppb. Defluorination of 6:2 FTS was calculated as follows:

$$\frac{\text{mol F release}}{\text{mol 6: 2 FTS}} = \frac{\text{ppb of F measured by probe}}{19\frac{g}{mol} \times \mu \text{M of 6: 2 FTS at time0}} \times 100\%$$

3.4. Assessment of desulfonation of 6:2 FTS by desulfonating enzymes

To examine if desulfonating enzymes are able to desulfonate 6:2 FTS, the taurine dioxygenase gene (tauD) and alkanesulfonate monooxygenase gene (ssuD and ssuE) genes of *Rhodococcus jostii* RHA1were identified from its whole genomic DNA in NCBI database. The genes encoding SsuD, SsuE, and TauD were overexpressed from pET11a in *E. coli* BL21(DE3) cells. Primers were designed based on DNA sequence of RHA1 published in NCBI. The successful amplification of ssuD, ssuE and tauD genes were checked by agarose gel. The PCR products were extracted using the QIAquick PCR Purification Kit for subsequent cloning.

3.4.1. Construction of expression plasmids

The cloning of tauD, ssuD and ssuE into expression vectors was performed in following steps. Gene tauD was PCR-amplified using the primers tauD-FW (5'-TAA GGA CAT ATG ATG AGC ACT GCA TTC GAA ACC AG-3') and tauD-RV (5'- TTA TTT GCT AGC CTA CTG CGG GCC GAC GG-3') which engineered to introduce restriction sites NdeI and NheI overhangs. In the second step, the T7 RNA polymerasedependent expression vector pET11a and PCR product of tauD were purified with the QIAquick PCR Purification Kit, and digested by restriction enzymes NdeI and NheI in 37°C water bath for 3 hours and 1 hour, respectively. The digest DNA was isolated by agarose gel running and purification. The resulting insert and truncated vector were fused by DNA ligation (Fig. 3.1). A negative control was set up for the background of self-ligating recipient plasmid backbone. One μL of ligation mixture was transformed
into NEB 5- α competent *E. coli* cells. The transformants were selected on LB (Luria-Bertani) medium agar plates containing 100 ppm ampicillin. Four colonies were selected on each plate and inoculated into 5 mL LB broth containing 100 ppm ampicillin, and incubated overnight at 37 °C with constant shaking (150 rpm). DNA was subsequently purified and diagnosed with restriction digest with enzymes NotI-HF and HpaI and agarose gel running.



Figure 3.1. Map of constructed recombinant plasmid pET11a-tauD.

The procedure for the construction of ssuD and ssuE expression plasmids were similar to the plasmid construction for tauD with the following exceptions (Fig. 3.2). The gene ssuD was PCR-amplified using the primers ssuD-FW (5'-TAA GGA CAT ATG ATG AGC ATC GAC TTC TAC TGG C-3') and ssuD-RV (5'- TTA TTT GCT AGC TCA CAG CAG GTC CTG TCC G-3'). The oligonucleotide primers used to PCRamplify ssuE gene were ssuE-FW (5'- TAA TTG CAT ATG ATG TCA CAG ACC AAC GTT CTC G-3') and ssuE-RV (5'- TTA TTT GCT AGC TCA GGC GTC GAC GAG CTG-3'). After transformation to NEB 5-α competent *E. coli* cells, the purified DNA of pET11a-ssuE plasmid was diagnosed with restriction digest with enzymes XhoI and HpaI.



Figure 3.2. Map of constructed recombinant plasmids pET11a-ssuD and pET11a-ssuE.

DNA vectors containing representative clones were submitted for sequence analysis at Eton Bioscience Sequencing. Confirmed that no mutations had been introduced into ssuD, ssuE, or tauD genes during PCR amplification or ligation, the expression plasmids were transformed to *E. coli* strain BL21(DE3) for subsequent overexpression, respectively.

3.4.2. Protein expression

For the production of TauD, a single colony of *E. coli* BL21(DE3) containing the appropriate expression plasmid was used to inoculate 10 mL LB media supplemented

with 100 ppm ampicillin, which was incubated overnight at 37 °C. A 20-25% inoculum of the 10 mL culture was used to inoculate 200 mL LB-Amp media, which was incubated 2 hours at 37 °C. To minimize the formation of inclusion bodies, which were observed when protein production was carried out at 30°C and 37 °C, cultures grown to an OD_{600} of 0.3-0.4 were cooled to 18 °C before induction. When the OD_{600} value reached 0.4-0.5, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.4 mM to induce the overexpression of TauD. The incubation was continued for 9.5 hours at 18 °C with constant shaking (150 rpm). Cells were harvested by centrifugation at 10,000 g for 10 min at 4 °C, washed in an excess of 25 mM Tris buffer, pH 8.0, and stored at -80°C as frozen pellets until further use. For the preparation of cell-free extracts, cells from 50 mL growth were resuspended in 2 mL TE buffer (25 mM Tris buffer, pH 8.0, and 1mM EDTA) containing 20% glycerol and sonicated. After centrifugation at 24,000 g for 1 hour, the supernatant was collected and stored at -80°C in 800 µL aliquots until further use. The procedure for the overexpression of SsuD and SsuE proteins were similar to the overexpression of TauD except that SsuD was induced at 4 °C for 3 days. The lysis buffer of SsuD and SsuE was 25mM potassium phosphate buffer (pH 7.5) with 10% glycerol.

3.4.3. Enzyme activity assay

The concentrations of total cellular proteins in the crude cell extracts were determined using a Pierce BCA protein assay kit (ThermoFisher Scientific, Waltham, MA) following the manufacturer's protocol. Taurine dioxygenase (TauD) activity assay was performed by using Ellman's reagent (5,5'-dithiobis(2-nitrobenzoic acid)), which produces a bright yellow color upon reaction with sulfite²⁵. Reaction was initiated by the addition of TauD crude enzyme to a reaction mixture containing 500 mM taurine, 1 mM α KG, 100 μ M FeCl₂ (freshly made), 200 μ M sodium ascorbate in 10 mM imidazole buffer, pH 6.9, at 30 °C for 3 minutes. The assay was stopped by the addition of 500 μ L sample to a cuvette containing 400 μ L urea (4.3 M) and 100 μ L Ellman's reagent (1 mg/mL in 100 mM sodium phosphate buffer, pH 7.0). The colorimetric reaction was allowed to develop at room temperature for 2-3 min and the absorbance was measured at 412 nm. A linear calibration curve of sulfite was developed from fresh sodium sulfite in deionized water, ranging from 50 to 500 μ M. An assay mixture without taurine was used as a blank. One unit of activity is defined as the amount of enzyme forming 1 μ mol of sulfite per minute at 30 °C.

Alkanesulfonate reductase (SsuE) activity assay was carried out by monitoring the absorbance decrease at 340 nm due to the oxidation of NADPH. The reaction was started by the addition of SsuE enzyme (0.01 μ M) to an assay mixture (1 mL) containing 500 μ M NADPH, 3 μ M Flavin mononucleotide (FMN) in 25 mM Tris buffer. An assay mixture without sulfonate was used as a blank. One unit of activity is defined as the amount of enzyme oxidizing 1 μ mol of NADPH per minute at 30 °C.

Alkanesulfonate monooxygenase (SsuD) activity was assayed with Ellman's reagent. The assay mixture (1.5 mL) contained 500 μ M NADPH, 3 μ M FMN, 500 μ M hexanesulfonic acid sodium salt, with SsuD to SsuE at a molar ratio of 1:3 in 25 mM potassium phosphate buffer. The reaction was initiated by the addition of SsuD to the

mixture and was stirred on a magnetic stirrer at 30 °C for 3 minutes. The sulfite production was determined after the addition of 100 μ L Ellman's reagent (1 mg/mL in 100 mM sodium phosphate buffer, pH 7.0). The colorimetric reaction was allowed to develop at room temperature for 2 min and the absorbance was measured at 412 nm. One unit of activity is defined as the amount of enzyme forming 1 μ mol of sulfite per minute at 30 °C.

3.4.4. Desulfonation of 6:2 FTS by overexpressed enzymes

Desulfonation of 6:2 FTS by overexpressed taurine dioxygenase (TauD) was assayed by determining the amount of sulfite release during 2 hours of incubation at 30 °C. The reaction mixture contained 500 μ M 6:2 FTS, 1 mM α KG, 100 μ M FeCl₂ (freshly made), 200 μ M sodium ascorbate in 10 mM imidazole buffer, pH 6.9. After 2-hour of incubation, the sulfite release was confirmed by the colorimetric assay as described above and quantified by using ion chromatography (Dionex IonPacTM AS19-4 μ m) (Thermo Scientific, Beverly, MA).

Desulfonation of 6:2 FTS by overexpressed alkanesulfonate reductase and monooxygenase (SsuE and SsuD) was carried out in a reaction mixture containing 500 μ M 6:2 FTS, 500 μ M NADPH, 3 μ M FMN, SsuD to SsuE at a ratio of 1:3 in 25 mM Tris buffer. The reaction mixture was stirred on a magnetic stirrer at 30 °C for 2 hours. The sulfite release was confirmed by the colorimetric assay as described above and quantified by using ion chromatography (Dionex IonPacTM AS19-4 μ m) (Thermo Scientific, Beverly, MA).

4. RESULTS AND DISCUSSION

4.1. Identification and characteristics of a 6:2 FTS-degrading strain, *Pseudomonas* SYC

A total of nine pure isolates were screened for the ability to grow on sulfur-free basal medium agar plate containing 6:2 FTS as the sole sulfur source. However, only three strains (named as No. 7, No. 11, and SYC) were able to utilize 6:2 FTS as the sole sulfur source in the liquid basal medium and produce significant biomass. Analysis of 16S rRNA gene sequences clearly demonstrated that isolates No. 7 and No. 11 are members of *Cupriavidus* species. SYC showed 96-99% sequence similarities to a number of *Pseudomonas* species. Out of the three isolates, neither *Cupriavidus* strain was capable of defluorinating 6:2 FTS after 10-day incubations at 30 °C at 150 rpm. It is likely due to both *Cupriavidus* strains lacking enzymes that can further transform and defluorinate 6:2 FTS. In contrast, isolate SYC exhibited high capability in defluorinating 6:2 FTS with nearly 0.9 mol F⁻ release / mol 6:2 FTS after 10-day incubation (Fig. 4.1). Therefore, isolate SYC was selected for further study on the basis of the extent of growth and defluorination potential, and designated strain *Pseudomonas* SYC.

Over 97% of 6:2 FTS was declined quickly after 24-hr incubation at 30 °C with a single colony inocula of *Pseudomonas* SYC in sulfur-free basal medium (Fig. 4.1). Meanwhile, the free fluoride concentration detected in the medium increased with time. This could be explained by the transformation of 6:2 FTS into less-fluorinated metabolites with fluoride release. The increase of fluoride concentration was observed

even though 6:2 FTS was no longer detected after day 3. This indicates a slow transformation of 6:2 FTS metabolites with fluoride being released. A control was also included using an uninoculated medium for which degradation was not observed within the same period of time.



Figure 4.1. Time course of biotransformation and defluorination of 6:2 FTS by the isolate strain *Pseudomonas* SYC. (A) 6:2 FTS biotransformation by strain SYC; 6:2 FTS was quantified with HPLC/MS, (B) 6:2 FTS defluorination by strain SYC; free fluoride was quantified with ion selective electrode.

A phylogenetic tree was constructed by alignment of 16S rRNA sequence of *Pseudomonas* SYC with strains having over 96% similarities based on the BLAST results. A relatively high bootstrap value (97%) linked strain *Pseudomonas* SYC to *Pseudomonas nitroreducens* NCBR 12694 (Fig. 4.2), indicating a strong phylogenetic relationship of *Pseudomonas* SYC to this strain.



Figure 4.2. A phylogenetic tree showing relative relationship between strain SYC and related *Pseudomonas* species. The phylogenetic tree was constructed using neighbor-joining method. Evolutionary distances were computed using the maximum composite likelihood method and are given in units of the number of base substitutions per site. Bootstrap support values from 1000 replicates are indicated at branch nodes. NCBI nucleotide accession numbers for each sequence are included in parentheses. An asterisk identifies the isolate 6:2 FTS-degrading strain.

Strain SYC has 97% similarity to a 6:2 FTOH degrader *Pseudomonas oleovorans*,³² indicating a substantial relationship between SYC and this 6:2 FTOHdegrading strain. In addition, SYC has 97% and 96% homology to *Pseudomonas aeruginosa* and *Pseudomonas putida*, respectively, which were capable of desulfurizing a broad range of aromatic sulfonates.^{34, 37} It is not surprising to observe the high homology of strain SYC to known desulfonating strains, since desulfonation is a presumable essential step in 6:2 FTS degradation. Interestingly, strains that have over 96% homology to SYC contain either alkanesulfonate monooxygenase gene (ssuD) or taurine dioxygenase gene (tauD) or both, suggesting that these two genes are likely to code enzymes responsible for the desulfonation of 6:2 FTS.

4.2. Factors affecting 6:2 FTS biodefluorination by pure strains

Since 6:2 FTS is structurally similar to 6:2 FTOH, experiments were conducted to determine if 6:2-FTOH degrading strains, *Rhodococcus jostii* RHA1 and *Pseudomonas oleovorans*, could degrade 6:2 FTS. As environment matrices contain a wide range of complex nutrients that are readily available for microbial growth, it is important to know whether 6:2 FTS-degrading strains would retain its ability to degrade 6:2 FTS when being provided with relatively more accessible sulfur source, such as sulfate.

To determine the effects of nutrients on 6:2 FTS defluorination by strain *Rhodococcus jostii* RHA1, *Pseudomonas oleovorans*, and isolate *Pseudomonas* SYC, the strains was initially grown in sulfate-contained medium without 6:2 FTS overnight as inoculum at time 0. When 6:2 FTS and sulfate were both present in the growth medium, all three strains lost their defluorination abilities toward 6:2 FTS. However, when 6:2 FTS was supplied as sole sulfur source with appropriate carbon source being provided, the strains were able to defluorinate 6:2 FTS starting as early as at the beginning of exponential phase.

37

4.2.1. Biodefluorination of 6:2 FTS by Rhodococcus jostii RHA1

As 6:2 FTS is a polyfluorinated compound, the amount of fluoride released from 6:2 FTS biodegradation was assessed for the degree of biodefluorination. Free fluoride was detected at early exponential phase (OD₆₀₀~0.2) with 0.15 mol F⁻ by ethanol-grown RHA1 and 0.07 mol F⁻ by butanol-grown RHA1 under sulfur free conditions (Fig. 4.3). With the generation of over 1.2 mol F⁻ out of 1 mol 6:2 FTS after 7-day incubation, it is apparent that 6:2 FTS was utilized as sole sulfur source and converted to transformed products by ethanol-grown or butanol-grown RHA1. However, no free fluoride was detected by taurine-grown RHA1 under sulfur-limiting condition. A possible reason is that ethanol and butanol could serve as inducers of defluorinating enzyme, presumably alcohol dehydrogenase (ADH)⁷⁶ and soluble butane monooxygenase (sBMO),⁵¹ respectively. In contrast, taurine was likely only served as carbon source, but not as an inducer of any defluorinating enzyme. Not surprisingly, there was no observation of fluoride release when sulfate was present in the medium, although RHA1 was incubated with ethanol as an inducer of defluorinating enzyme. These results indicate that 6:2 FTS needs to be desulfonated first for further defluorination to occur. The presence of sulfate in the medium might repress the expression of enzymes for sulfur removal from 6:2 FTS, which is in consistent with the characteristics of reported desulfonating enzymes: taurine dioxygenase (TauD) and two-component alkanesulfonate monooxygenase system (SsuE/D).



Figure 4.3. *Rhodococcus jostii* RHA1 growth curve and time course of 6:2 FTS defluorination. (A) growth curve of RHA1 with different substrates (measured with spectrophotometer), (B) fluoride released by RHA1 under different growth conditions (measured with ion selective electrode).

Even though RHA1 can defluorinate 6:2 FTS, the bacterium cannot grow on 6:2 FTS as the sole carbon source. Currently, no information is available regarding 6:2 FTS biotransformation pathways and potential transformation products. Based on the amounts of fluoride released, 6:2 FTS was biodefluorinated extensively under sulfurlimiting conditions when ethanol or butanol supplied as carbon source as well as 6:2 FTS served as the sole sulfur source.

4.2.2. Biodefluorination of 6:2 FTS by Pseudomonas oleovorans

Pseudomonas oleovoarns showed ability to defluorinate 6:2 FTS in the presence of ethanol and n-octane when 6:2 FTS was provided as sole sulfur source. Around 0.08 mol of free fluoride was detected at early exponential phase (OD_{600} ~0.14) by ethanolgrown *P. ole* under sulfur free condition (Fig. 4.4). In contrast, there was no observation of fluoride release by octane-grown *P. ole* at either early exponential phase or the middle of exponential phase. At stationary phase, ethanol-grown *P. ole* defluorinated approximately 0.9 mol F⁻ and octane-grown *P. ole* defluorinated around 0.8 mol F⁻. The reasons why there were differences at the time point for defluorination start to occur by ethanol-grown *P. ole* and octane-grown *P. ole* were unclear. It's possible that the supplies of ethanol and n-octane induced expressions of different enzymes, presumably alcohol dehydrogenase (ADH) and alkane monooxygenase,³³ respectively. As a result, the different downstream biotransformation mechanisms might be performed. Although containing tauD gene, *P. ole* didn't show significantly active growth when taurine was provided as sole carbon source. It might be due to lack of enzymes that could subsequently transform aminoacetaldehyde (converted from taurine after desulfonation) to an available form of carbon source for *P. ole* to utilize.



Figure 4.4. *Pseudomonas oleovorans* growth curve and time course of 6:2 FTS defluorination. (A) growth curve of *P. ole* with different substrates (measured with spectrophotometer), (B) fluoride released by *P. ole* under different growth conditions (measured with ion selective electrode).

4.2.3. Biodefluorination of 6:2 FTS by Pseudomonas SYC

The biodefluorination by *Pseudomonas* SYC under different growth conditions was similar to that of RHA1 but with lower extent. Free fluoride was detected at early exponential phase (OD_{600} ~0.1) with 0.09 mol F⁻ by ethanol-grown SYC under sulfurlimiting condition (Fig. 4.5). Ethanol-grown and butanol-grown SYC released approximately 0.7 mol F⁻ out of 1 mol 6:2 FTS after 7-day incubation. No free fluoride was detected by taurine-grown SYC under sulfur-free condition, nor by ethanol-grown SYC in the presence of sulfate in the medium. The reasons of different levels performed in the amount of fluoride released by strain RHA1 and SYC were not known. It is possible that these strains have different susceptibility to biotransformed product toxicity of 6:2 FTS.



Figure 4.5. *Pseudomonas* SYC growth curve and time course of 6:2 FTS defluorination. (A) growth curve of SYC with different substrates (measured with spectrophotometer), (B) fluoride released by SYC under different growth conditions (measured with ion selective electrode).

4.3. Desulfonation of 6:2 FTS by enzymes

4.3.1. Relationships of putative desulfonating bacteria

The sequences of sulfur-specific metabolic genes (ssuD and tauD) of *Rhodococcus jostii* RHA1, *Pseudomonas oleovorans*, and putative desulfonating bacteria were searched in the NCBI database using the BLAST program and aligned using the ClustalW program.

A neighbor-joining tree of ssuD genes (Fig. 4.6) showed that the ssuD sequence of *P. ole* tended to cluster with that of *Gordonia* sp. NB4-1Y, a reported 6:2 FTS degrading strain. The ssuD gene of RHA1 has relatively close relationship with that of NB4-1Y compared to the others.



Figure 4.6. Neighbor-joining tree illustrating the relationships of ssuD genes of putative desulfonating bacteria. Evolutionary distances were computed using the maximum composite likelihood method and are given in units of the number of base substitutions per site. Bootstrap support is shown at each node as a percentage, based on 1000 resampled datasets. NCBI nucleotide accession numbers for each sequence are included in parentheses.

Similarly, the neighbor-joining tree of tauD gene was also generated (Fig. 4.7). The tauD gene of *P. ole* has relatively close relationship with that of NB4-1Y. In contrast, the relationship of tauD genes between RHA1 and NB4-1Y is relatively distanced.



Figure 4.7. Neighbor-joining tree illustrating the relationships of tauD genes of putative desulfonating bacteria. Evolutionary distances were computed using the maximum composite likelihood method and are given in units of the number of base substitutions per site. Bootstrap support is shown at each node as a percentage, based on 1000 resampled datasets. NCBI nucleotide accession numbers for each sequence are included in parentheses.

The strains selected for the generation of ssuD gene and tauD gene trees are either reported to have desulfurized sulfonates, or strains that have 99% homology to the isolate SYC. *Pseudomonas aeruginosa* PAO1 and *Pseudomonas putida* S-313 are known of ability to desulfurize a wide range of aromatic sulfonates. The organosulfur metabolism in *Escherichia coli* and *Bacillus subtilis* have been well studied. *Klebsiella oxytoca* is a SDS-degrading strain isolated from soil. *Corynebacterium glutamicum* is able to utilized a number of sulfonates and sulfonate esters as sulfur sources. The close relationships of ssuD and tauD genes of RHA1, *P. ole*, and strains with 99% homology to SYC with those of the reported desulfonating strains indicated that ssuD and tauD might play important roles in the desulfonation of 6:2 FTS.

4.3.2. Cloning and overexpression of desulfonating enzymes in *E. coli*

ssuD, ssuE and tauD genes were cloned and expressed for the examine of the capability of 6:2 FTS desulfonation. The amplifications of target genes were confirmed by running agarose gel. The band of 1.2 kilobase pairs (kb) for ssuD, 0.6 kb for ssuE, and 0.9 kb for tauD were observed (Fig. A1 in appendix). The digested unique DNA fragment size (listed in Table. A1 in appendix) on the agarose gel (Fig. A2 in appendix) indicated the successful ligation of each target gene with pET11a vectors, which was further confirmed by the sequencing of each constructed plasmid (Fig. A3-1, Fig. A3-2, and Fig. A3-3 in appendix). The three proteins of interest were successfully overexpressed in E. coil BL21 (DE3) and confirmed by running SDS-PAGE (Fig.4.8). The calculated subunit molecular mass from the tauD gene sequence of 31.8 kilodalton (kDa) was approximately estimated by SDS-PAGE analysis as 35 kDa. The subunit molecular masses calculated from the SsuE and SsuD amino acid sequence of 19.4 kDa and 41.8 kDa, were estimated by SDS-PAGE analysis as around 22.5 kDa and 38 kDa, respectively. The subtle deviations between theoretically predicted and SDS PAGEdisplayed molecular weights might be due to the high content of acidic amino acids in the proteins.⁷⁷



Figure 4.8. Overexpressed proteins of interest (SsuE, SsuD, and TauD).

4.3.3. Activities of enzymes

The concentrations of expressed enzymes in the *E. coli* were as follows: TauD 0.72 mg/mL, SsuE 1.93 mg/mL, SsuD 0.85 mg/mL. The activity of TauD in crude cell extracts was examined for taurine desulfonation by following taurine-dependent sulfite release. Sulfite release was not detected in cell extracts prepared from uninduced cells control or in crude extracts from induced cells incubated without taurine. TauD in crude cell extracts showed a taurine desulfonating activity of 0.75 U/mg of protein. SsuE in crude cell extracts showed a NADPH oxidizing activity of 0.3 U/mg of SsuE protein. The activity of SsuD in crude cell extracts was examined for sodium hexanesulfonate desulfonating activity of 0.047 U/mg of SsuD protein.

4.3.4. Desulfonation of 6:2 FTS in crude cell extracts

Desulfonation of 6:2 FTS by TauD was evaluated by determining the amount of sulfite formed during 2 hours of incubation at 30 °C in 10 mM imidazole buffer containing 2 μ M TauD, 500 μ M 6:2 FTS, 1 mM α KG, 100 μ M FeCl₂, and 200 μ M sodium ascorbate. Sulfite release was not detected in cell extracts prepared from uninduced cells or in extracts from induced cells incubated without 6:2 FTS. Around 81.64 μ M of sulfite and 19.15 μ M of sulfate, which might be oxidized from sulfite, were detected in on chromatography, indicating that TauD is capable of desulfonating 6:2 FTS.

Desulfonation of 6:2 FTS by SsuE/D was evaluated by determining the amount of sulfite formed during 2 hours of incubation at 30 °C in 25 mM Tris buffer containing 500 μ M 6:2 FTS, 500 μ M NADPH, 3 μ M FMN, SsuD to SsuE at a molar ratio of 1:3. Around 10.65 μ M of sulfite was detected in ion chromatography, indicating that SsuD/E system is capable of desulfonating 6:2 FTS.

5. SUMMARY, CONCLUSIONS AND FUTURE STUDIES

5.1. Summary and conclusions

6:2 Fluorotelomer sulfonate (6:2 FTS) is often utilized as mist suppressing agent in non-decorative hard chrome plating industry as an alternative to a banned chemical perfluorooctane sulfonate (PFOS). As a transformation metabolite from constituents in AFFFs, not surprisingly, frequent detections of 6:2 FTS with high concentrations were found in groundwater and soil that are contaminated with AFFFs. While biodegradation of 6:2 FTS have been suggested in activated sludge, aerobic river sediments, and pure culture studies, the conditions caused biodegradation of 6:2 FTS are unclear and have not been fully investigated. In addition, little was known about the enzymes responsible for the desulfonation of 6:2 FTS.

In this study, three 6:2 FTS-degrading strains were identified. A rhizosphere soil isolate, *Pseudomonas* SYC, was able to utilize 6:2 FTS as sole sulfur source and biotransform nearly 97% of 6:2 FTS with free fluoride release within 7 days of incubation. Based on the BLAST result of 16S rRNA gene sequence alignment, the isolate strain SYC has over 96% homology to known desulfonating strains. Interestingly, all strains that have over 96% homology to SYC contain either alkanesulfonate monooxygenase gene (ssuD) or taurine dioxygenase gene (tauD) or both, suggesting that these two genes are likely to code enzymes responsible for the desulfonation of 6:2 FTS.

Two 6:2 FTOH-degrading strains, *Rhodococcus jostii* RHA1 and *Pseudomonas oleovorans*, also showed ability to defluorinate 6:2 FTS under sulfur-limiting conditions.

Ethanol-grown and n-octane-grown *P. ole* were able to defluorinate 0.8 to 0.9 mol F⁻/mol 6:2 FTS on day 7. Relatively higher concentration of fluoride was released by ethanol-grown and butanol-grown RHA1, nearly 1.4 mol F⁻/mol 6:2 FTS on day 7. Both 6:2 FTOH-degrading strains contained ssuD gene and tauD gene.

None of the three strains were able to defluorinate 6:2 FTS when taurine was supplied as sole carbon source, suggesting that necessary inducers of defluorinating enzymes are required for the defluorination to occur, such as ethanol, 1-butanol, and n-octane. The 6:2 FTS defluorination ability of these three strains can be retained when the growth medium is in absence of alternative sulfur source, sulfate. The phenomenon of losing 6:2 FTS biodefluorination ability in the presence of sulfate indicated that the expression of desulfonating enzymes involved in desulfonation of 6:2 FTS was likely repressed, suggesting that desulfonation is the first step in 6:2 FTS metabolism by 6:2 FTS-degrading strains.

This was the first study using in vitro enzymatic assays to demonstrate that two known desulfonating enzyme systems, two-component alkanesulfonate monooxygenase system (SsuE/D) and taurine dioxygenase (TauD), play a role in desulfonation of 6:2 FTS. In this study, the genes of ssuD, ssuE and tauD were successfully cloned and expressed from plasmid pET11a in *E. coli* BL21(DE3). The ability of 6:2 FTS desulfonation by these two enzyme systems was confirmed by detections of sulfite release in liquid medium, providing fundamental knowledge for future studies on the molecular biology of 6:2 FTS metabolism in bacteria.

48

5.2. Future studies

This study demonstrated that the two-component alkanesulfonate monooxygenase system (SsuE/D) and taurine dioxygenase (TauD) are responsible for desulfonation of 6:2 FTS. The defluorination of 6:2 FTS by three pure cultures were also discussed. However, the results of this study brought some questions to several future studies. Below are suggestions for future studies that might answer the remained questions in this study.

Two *Cupriavidus* strains were isolated from rhizosphere in this study. They were capable of utilizing 6:2 FTS as sole sulfur source but not able to defluorinate it. It is likely that these two *Cupriavidus* strains lack enzymes to further biotransform 6:2 FTS. Thus a 6:2 FTS-degrading consortia can be developed by using the 6:2 FTS-utilizing strains as well as 6:2 FTOH-degrading strains to enhance the biodegradation of 6:2 FTS.

Rapid 6:2 FTS biotransformation with a half-life of around 5 days was reported in aerobic river sediment.²¹ The rapid transformation suggested that 6:2 FTS desulfonation occurred although there was sulfate present (~0.3mM). It is possible that the desulfonating enzymes were induced after the sulfate were consumed. This aspect was not explored in this study. A better understanding of the sulfate concentration affecting 6:2 FTS biotransformation potential and its degradation rate is also needed. Accordingly, it is necessary to use artificial microbial community or enrichment cultures to evaluate their desulfonation ability in the presence of different level of sulfate.

The synthesis of SsuE, SsuD, and TauD was expressed during growth with alkanesulfonates under sulfur-limiting condition, but repressed in the presence of

inorganic sulfur source. However, in the environment, the concentrations of sulfate, sulfite, sulfide, and cysteine can be significant. Therefore, the regulations of ssuE, ssuD, and tauD gene expressions in 6:2 FTS-degrading strains under sulfur-contained conditions are needed for the enhanced desulfonation of 6:2 FTS, even though when inorganic sulfur sources are present.

Further studies are also needed to identify other desulfonating enzymes that can desulfonate 6:2 FTS under different redox conditions.

REFERENCES

1. Yang, X.; Huang, J.; Zhang, K.; Yu, G.; Deng, S.; Wang, B., Stability of 6:2 fluorotelomer sulfonate in advanced oxidation processes: degradation kinetics and pathway. *Environmental Science and Pollution Research* **2014**, *21*, (6), 4634-4642.

 Field, J. A.; Seow, J., Properties, occurrence, and fate of fluorotelomer sulfonates. *Critical Reviews in Environmental Science and Technology* 2017, *47*, (8), 643-691.

3. Sheng, N.; Zhou, X.; Zheng, F.; Pan, Y.; Guo, X.; Guo, Y.; Sun, Y.; Dai, J., Comparative hepatotoxicity of 6:2 fluorotelomer carboxylic acid and 6:2 fluorotelomer sulfonic acid, two fluorinated alternatives to long-chain perfluoroalkyl acids, on adult male mice. *Archives of Toxicology* **2017**, *91*, (8), 2909-2919.

4. Jin, L.; Jiang, C.; Zhang, P., Photochemical decomposition of 1H,1H,2H,2Hperfluorooctane sulfonate (6:2 FTS) induced by ferric ions. *Journal of Environmental Sciences* **2017**, *51*, 120-127.

Poulsen, P. B.; Gram, L. K.; Jensen, A. A.; Rasmussen, A. A.; Ravn, C.; Møller,
 P.; Jørgensen, C. R.; Løkkegaard, K., Substitution of PFOS for use in non-decorative
 hard chrome plating. *Danish Ministry of the Environment* 2011, *1371*, 1-77.

6. Buck, R. C.; Franklin, J.; Berger, U.; Conder, J. M.; Cousins, I. T.; de Voogt, P.; Jensen, A. A.; Kannan, K.; Mabury, S. A.; van Leeuwen, S. P., Perfluoroalkyl and polyfluoroalkyl substances in the environment: terminology, classification, and origins. *Integrated Environmental Assessment and Management* **2011**, *7*, (4), 513-541. Place, B. J.; Field, J. A., Identification of novel fluorochemicals in aqueous filmforming foams used by the US military. *Environmental Science & Technology* 2012, *46*, (13), 7120-7127.

 Dauchy, X.; Boiteux, V.; Bach, C.; Colin, A.; Hemard, J.; Rosin, C.; Munoz, J.
 F., Mass flows and fate of per- and polyfluoroalkyl substances (PFASs) in the wastewater treatment plant of a fluorochemical manufacturing facility. *Science of the Total Environment* 2017, *576*, 549-558.

9. Fang, C.; Megharaj, M.; Naidu, R., Chemical oxidization of some AFFFs leads to the formation of 6:2 FTS and 8:2 FTS. *Environmental Toxicology and Chemistry* **2015**, *34*, (11), 2625-2628.

10. Harding-Marjanovic, K. C.; Houtz, E. F.; Yi, S.; Field, J. A.; Sedlak, D. L.; Alvarez-Cohen, L., Aerobic biotransformation of fluorotelomer thioether amido sulfonate (Lodyne) in AFFF-amended microcosms. *Environmental Science & Technology* **2015**, *49*, (13), 7666-7674.

11. Kärrman, A.; Elgh-Dalgren, K.; Lafossas, C.; Møskeland, T., Environmental levels and distribution of structural isomers of perfluoroalkyl acids after aqueous fire-fighting foam (AFFF) contamination. *Environmental Chemistry* **2011**, *8*, (4), 372-380.

12. Schultz, M. M.; Barofsky, D. F.; Field, J. A., Quantatitive determination of fluorotelomer sulfonates in groundwater by LC/MS/MS. *Environmental Science & Technology* **2004**, *38*, 1828-1835.

Backe, W. J.; Day, T. C.; Field, J. A., Zwitterionic, cationic, and anionic
 fluorinated chemicals in aqueous film forming foam formulations and groundwater from

U.S. military bases by nonaqueous large-volume injection HPLC-MS/MS.

Environmental Science & Technology 2013, 47, (10), 5226-5234.

14. Willach, S.; Brauch, H. J.; Lange, F. T., Contribution of selected perfluoroalkyl and polyfluoroalkyl substances to the adsorbable organically bound fluorine in German rivers and in a highly contaminated groundwater. *Chemosphere* **2016**, *145*, 342-350.

15. Boiteux, V.; Dauchy, X.; Bach, C.; Colin, A.; Hemard, J.; Sagres, V.; Rosin, C.; Munoz, J. F., Concentrations and patterns of perfluoroalkyl and polyfluoroalkyl substances in a river and three drinking water treatment plants near and far from a major production source. *Science of the Total Environment* **2017**, *583*, 393-400.

16. Fernandez, N. A.; Rodriguez-Freire, L.; Keswani, M.; Sierra-Alvarez, R., Effect of chemical structure on the sonochemical degradation of perfluoroalkyl and polyfluoroalkyl substances (PFASs). *Environmental Science: Water Research & Technology* **2016**, *2*, (6), 975-983.

17. Fath, A.; Sacher, F.; McCaskie, J. E., Electrochemical decomposition of fluorinated wetting agents in plating industry waste water. *Water Science & Technology* 2016, *73*, (7), 1659-1666.

18. Zhuo, Q.; Li, X.; Yan, F.; Yang, B.; Deng, S.; Huang, J.; Yu, G., Electrochemical oxidation of 1H,1H,2H,2H-perfluorooctane sulfonic acid (6:2 FTS) on DSA electrode: operating parameters and mechanism. *Journal of Environmental Sciences* **2014**, *26*, (8), 1733-1739.

19. Park, S.; Lee, L. S.; Medina, V. F.; Zull, A.; Waisner, S., Heat-activated persulfate oxidation of PFOA, 6:2 fluorotelomer sulfonate, and PFOS under conditions suitable for in-situ groundwater remediation. *Chemosphere* **2016**, *145*, 376-383.

20. Wang, N.; Liu, J.; Buck, R. C.; Korzeniowski, S. H.; Wolstenholme, B. W.; Folsom, P. W.; Sulecki, L. M., 6:2 fluorotelomer sulfonate aerobic biotransformation in activated sludge of waste water treatment plants. *Chemosphere* **2011**, *82*, (6), 853-858.

Zhang, S.; Lu, X.; Wang, N.; Buck, R. C., Biotransformation potential of 6:2
fluorotelomer sulfonate (6:2 FTSA) in aerobic and anaerobic sediment. *Chemosphere*2016, *154*, 224-230.

22. Van Hamme, J. D.; Bottos, E. M.; Bilbey, N. J.; Brewer, S. E., Genomic and proteomic characterization of *Gordonia* sp. NB4-1Y in relation to 6:2 fluorotelomer sulfonate biodegradation. *Microbiology* **2013**, *159*, (8), 1618-1628.

23. Shaw, D. M. J.; Munoz, G.; Bottos, E. M.; Duy, S. V.; Sauve, S.; Liu, J.; Van Hamme, J. D., Degradation and defluorination of 6:2 fluorotelomer sulfonamidoalkyl betaine and 6:2 fluorotelomer sulfonate by *Gordonia* sp. strain NB4-1Y under sulfur-limiting conditions. *Science of the Total Environment* **2018**, *647*, 690-698.

24. Key, B. D.; Howell, R. D.; Criddle, C. S., Defluorination of organofluorine sulfur compounds by *Pseudomonas* Sp. strain D2. *Environmental Science & Technology* **1998**, *32*, (15), 2283-2287.

25. Eichhorn, E.; van der Ploeg, J. R.; Kertesz, M. A.; Leisinger, T., Characterization of α-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *The Journal of Biological Chemistry* **1997**, *272*, (37), 23031-23036.

26. Gahan, J.; Schmalenberger, A., The role of bacteria and mycorrhiza in plant sulfur supply. *Frontiers in Plant Science* **2014**, *5*, 7231-7237.

27. Lu, H.; Sun, J.; Zhu, L., The role of artificial root exudate components in facilitating the degradation of pyrene in soil. *Scientific Reports* **2017**, *7*, (7130), 1-10.

 Shukor, M. Y.; Husin, W. S. W.; Rahman, M. F. A.; Shamaan, N. A.; Syed, M.
 A., Isolation and characterization of an SDS-degrading *Klebsiella oxytoca. Journal of Environmental Biology* 2009, *30*, (1), 129-134.

29. Shahbazi, R.; Kasra-Kermanshahi, R.; Gharavi, S.; Moosavi-Nejad, Z.;

Borzooee, F., Screening of SDS-degrading bacteria from car wash wastewater and study of the alkylsulfatase enzyme activity. *Iranian Journal of Microbiology* **2013**, *5*, 153-158.

30. Liu, J.; Wang, N.; Szostek, B.; Buck, R. C.; Panciroli, P. K.; Folsom, P. W.; Sulecki, L. M.; Bellin, C. A., 6-2 Fluorotelomer alcohol aerobic biodegradation in soil and mixed bacterial culture. *Chemosphere* **2010**, *78*, (4), 437-444.

31. Zhao, L.; Folsom, P. W.; Wolstenholme, B. W.; Sun, H.; Wang, N.; Buck, R. C.,
6:2 fluorotelomer alcohol biotransformation in an aerobic river sediment system. *Chemosphere* 2013, *90*, (2), 203-209.

32. Kim, M. H.; Wang, N.; McDonald, T.; Chu, K. H., Biodefluorination and biotransformation of fluorotelomer alcohols by two alkane-degrading *Pseudomonas* strains. *Biotechnology and Bioengineering* **2012**, *109*, (12), 3041-3048.

33. Kim, M. H.; Wang, N.; Chu, K. H., 6:2 Fluorotelomer alcohol (6:2 FTOH)
biodegradation by multiple microbial species under different physiological conditions. *Applied Microbiology & Biotechnology* 2013, 98, 1831-1840.

34. Kertesz, M. A., Riding the sulfur cycle - metabolism of suldonates and sulfate esters in Gram-negative bacteria. *FEMS Microbiology Reviews* **1999**, *24*, 135-175.

35. van der Ploeg, J.; Eichhorn, E.; Leisinger, T., Sulfonate-sulfur metabolism and its regulation in *Escherichia coli*. *Archives of Microbiology* **2001**, *176*, (1-2), 1-8.

36. Gao, B.; Ellis, H. R., Altered mechanism of the alkanesulfonate FMN reductase with the monooxygenase enzyme. *Biochemical and Biophysical Research Communications* **2005**, *331*, (4), 1137-1145.

37. Vermeij, P.; Wietek, C.; Kahnert, A.; WuÈest, T.; Kertesz, M. A., Genetic organization of sulphur-controlled aryl desulphonation in *Pseudomonas putida* S-313.

Molecular Microbiology 1999, 32, (5), 913-926.

38. Wei, C.; Wang, Q.; Song, X.; Chen, X.; Fan, R.; Ding, D.; Liu, Y., Distribution, source identification and health risk assessment of PFASs and two PFOS alternatives in groundwater from non-industrial areas. *Ecotoxicology and Environmental Safety* **2018**, *152*, 141-150.

39. Hansen, K. J.; Clemen, L. A.; Ellefson, M. E.; Johnson, H. O., Compoundspecific, quantitative characterization of organic fluorochemicals in biological matrices. *Environmental Science & Technology* **2001**, *35*, (4), 766-770.

Martin, J. W.; Kannan, K.; Berger, U.; Voogt, P. D.; Field, J.; Franklin, J.; Giesy,
J. P.; Harner, T.; Muir, D. C. G.; Scott, B.; Kaiser, M.; Järnberg, U.; Jones, K. C.;
Mabury, S. A.; Schroeder, H.; Simcik, M.; Sottani, C.; Bavel, B. V.; Kärrman, A.;
Lindström, G.; Leeuwen, S. V., Analytical challenges hamper perfluoroalkyl research. *Environmental Science & Technology* 2004, *38*, (13), 248A-255A.

41. Barzen-Hanson, K. A.; Davis, S. E.; Kleber, M.; Field, J. A., Sorption of fluorotelomer sulfonates, fluorotelomer sulfonamido betaines, and a fluorotelomer sulfonamido amine in national foam aqueous film-forming foam to soil. *Environmental Science and Technology* **2017**, *51*, (21), 12394-12404.

42. Wang, T.; Vestergren, R.; Herzke, D.; Yu, J.; Cousins, I. T., Levels, isomer profiles, and estimated riverine mass discharges of perfluoroalkyl acids and fluorinated alternatives at the mouths of Chinese rivers. *Environmental Science and Technology* **2016**, *50*, (21), 11584-11592.

43. Houtz, E. F.; Sutton, R.; Park, J. S.; Sedlak, M., Poly- and perfluoroalkyl substances in wastewater: significance of unknown precursors, manufacturing shifts, and likely AFFF impacts. *Water Research* **2016**, *95*, 142-149.

44. Ruan, T.; Lin, Y.; Wang, T.; Liu, R.; Jiang, G., Identification of novel polyfluorinated ether sulfonates as PFOS alternatives in municipal sewage sludge in China. *Environmental Science & Technology* **2015**, *49*, (11), 6519-6527.

45. Ahrens, L.; Felizeter, S.; Sturm, R.; Xie, Z.; Ebinghaus, R., Polyfluorinated compounds in waste water treatment plant effluents and surface waters along the River Elbe, Germany. *Marine Pollution Bulletin* **2009**, *58*, (9), 1326-1333.

46. Eggen, T.; Moeder, M.; Arukwe, A., Municipal landfill leachates: a significant source for new and emerging pollutants. *Science of the Total Environment* 2010, *408*, (21), 5147-5157.

47. Lee, H.; Mabury, S. A., A pilot survey of legacy and current commercial fluorinated chemicals in human sera from United States donors in 2009. *Environmental Science & Technology* **2011**, *45*, (19), 8067-8074.

48. Ahrens, L.; Norstrom, K.; Viktor, T.; Cousins, A. P.; Josefsson, S., Stockholm Arlanda Airport as a source of per- and polyfluoroalkyl substances to water, sediment and fish. *Chemosphere* **2015**, *129*, 33-38.

49. Lee, D. G.; Chu, K. H., Effects of growth substrate on triclosan biodegradation potential of oxygenase-expressing bacteria. *Chemosphere* **2013**, *93*, (9), 1904-1911.

Masai, E.; Yamada, A.; Healy, J. M.; Hatta, T.; Kimbara, K.; Fukuda, M.; Yano,
 K., Characterization of biphenyl catabolic genes of gram-positive polychlorinated
 biphenyl degrader *Rhodococcus* sp. Strain RHA1. *Applied and Environmental Microbiology* 1995, *61*, (6), 2079-2085.

51. Hand, S.; Wang, B.; Chu, K. H., Biodegradation of 1,4-dioxane: effects of enzyme inducers and trichloroethylene. *Science of the Total Environment* **2015**, *520*, 154-159.

52. McLeod, M. P.; Warren, R. L.; Hsiao, W. W.; Araki, N.; Myhre, M.; Fernandes,C.; Miyazawa, D.; Wong, W.; Lillquist, A. L.; Wang, D.; Dosanjh, M.; Hara, H.;

Petrescu, A.; Morin, R. D.; Yang, G.; Stott, J. M.; Schein, J. E.; Shin, H.; Smailus, D.;

Siddiqui, A. S.; Marra, M. A.; Jones, S. J.; Holt, R.; Brinkman, F. S.; Miyauchi, K.;

Fukuda, M.; Davies, J. E.; Mohn, W. W.; Eltis, L. D., The complete genome of

Rhodococcus sp. RHA1 provides insights into a catabolic powerhouse. Proceedings of

the National Academy of Sciences of the United States of America 2006, 103, (42), 15582-15587.

53. Schmalenberger, A.; Hodge, S.; Hawkesford, M. J.; Kertesz, M. A., Sulfonate desulfurization in *Rhodococcus* from wheat rhizosphere communities. *FEMS Microbiology Ecology* **2009**, *67*, (1), 140-150.

54. Baptist, J. N.; Gholson, R. K.; Coon, M. J., Hydrocarbon oxidation by a bacterial enzyme system: I. products of octane oxidation. *Biochimica et Biophysica Acta* **1963**, *69*, 40-47.

55. Chen, Q.; Janssen, D. B.; Witholt, B., Growth on octane alters the membrane lipid fatty acids of *Pseudomonas oleovorans* due to the induction of alkB and synthesis of octanol. *Journal of Bacteriology* **1995**, *177*, (23), 6894-6901.

56. van Beilen, J. B.; Kingma, J.; Witholt, B., Substrate specificity of the alkane hydroxylase system of *Pseudomonas oleovorans* GPo1. *Enzyme and Microbial Technology* **1994**, *16*, 904-911.

57. Xu, J.; Qiu, X.; Dai, J.; Cao, H.; Yang, M.; Zhang, J.; Xu, M., Isolation and characterization of a *Pseudomonas oleovorans* degrading the chloroacetamide herbicide acetochlor. *Biodegradation* **2006**, *17*, (3), 219-225.

58. Endoh, T.; Kasuga, K.; Horinouchi, M.; Yoshida, T.; Habe, H.; Nojiri, H.; Omori, T., Characterization and identification of genes essential for dimethyl sulfide utilization in *Pseudomonas putida* strain DS1. *Applied Microbiology and Biotechnology* **2003**, *62*, (1), 83-91. 59. Quadroni, M.; Staudenmann, W.; Kertesz, M.; James, P., Analysis of global responses by protein and peptide fingerprinting of proteins isolated by two-dimensional gel electrophoresis. *European Journal of Biochemistry* **1996**, *239*, 773-781.

Kertesz, M. A.; Wietek, C., Desulfurization and desulfonation: applications of sulfur-controlled gene expression in bacteria. *Applied Microbiology and Biotechnology* 2001, *57*, (4), 460-466.

61. Cook, A. M.; Laue, H.; Junker, F., Microbial desulfonation. *FEMS Microbiology Reviews* **1999**, *22*, 399-419.

Kertesz, M. A.; Schmidt-Larbig, K.; Wüest, T., A novel reduced flavin mononucleotide-dependent methanesulfonate sulfonatase encoded by the sulfur-regulated msu operon of *Pseudomonas aeruginosa*. *Journal of Bacteriology* 1999, *181*, (5), 1464-1473.

63. Omori, T.; Saiki, Y.; Kasuga, K.; Kodama, T., Desulfurization of alkyl and aromatic sulfides and sulfonates by dibenzothiophene-desulfurizing *Rhodococcus* sp. Strain SY1. *Bioscience, Biotechnology, and Biochemistry* **2014**, *59*, (7), 1195-1198.

64. Ellis, H. R., Mechanism for sulfur acquisition by the alkanesulfonate monooxygenase system. *Bioorganic Chemistry* **2011**, *39*, (5-6), 178-184.

65. Henderson, K. L.; Li, M.; Martinez, S.; Lewis, E. A.; Hausinger, R. P.; Emerson,
J. P., Global stability of an α-ketoglutarate-dependent dioxygenase (TauD) and its
related complexes. *Biochimica et Biophysica Acta* 2017, *1861*, (5 Pt A), 987-994.

66. Elkins, J. M.; Ryle, M. J.; Clifton, I. J.; Dunning Hotopp, J. C.; Lloyd, J. S.; Burzlaff, N. I.; Baldwin, J. E.; Hausinger, R. P.; Roach, P. L., X-ray crystal structure of *Escherichia coli* taurine/α-ketoglutarate dioxygenase complexed to ferrous iron and substrates. *Biochemistry* **2002**, *41*, 5185-5192.

67. Ryle, M. J.; Padmakumar, R.; Hausinger, R. P., Stopped-flow kinetic analysis of *Escherichia coli* taurine/α-ketoglutarate dioxygenase: interactions with α-ketoglutarate, taurine, and oxygen. *Biochemistry* **1999**, *38*, 15278-15286.

68. Robbins, J. M.; Ellis, H. R., Steady-state kinetic isotope effects support a complex role of Arg226 in the proposed desulfonation mechanism of alkanesulfonate monooxygenase. *Biochemistry* **2014**, *53*, (1), 161-168.

69. Driggers, C. M.; Dayal, P. V.; Ellis, H. R.; Karplus, P. A., Crystal structure of *Escherichia coli* SsuE: defining a general catalytic cycle for FMN reductases of the flavodoxin-like superfamily. *Biochemistry* **2014**, *53*, (21), 3509-3519.

70. Armacost, K.; Musila, J.; Gathiaka, S.; Ellis, H. R.; Acevedo, O., Exploring the catalytic mechanism of alkanesulfonate monooxygenase using molecular dynamics. *Biochemistry* **2014**, *53*, (20), 3308-3317.

Xiong, J.; Ellis, H. R., Deletional studies to investigate the functional role of a dynamic loop region of alkanesulfonate monooxygenase. *Biochimica et Biophysica Acta* 2012, *1824*, (7), 898-906.

72. Koch, D. J.; Ruckert, C.; Rey, D. A.; Mix, A.; Puhler, A.; Kalinowski, J., Role of the ssu and seu genes of *Corynebacterium glutamicum* ATCC 13032 in utilization of sulfonates and sulfonate esters as sulfur sources. *Applied and Environmental Microbiology* **2005**, *71*, (10), 6104-6114.

73. Eichhorn, E.; van der Ploeg, J. R.; Leisinger, T., Characterization of a twocomponent alkanesulfonate monooxygenase from *Escherichia coli*. *The Journal of Biological Chemistry* **1999**, *274*, (38), 26639-26646.

74. Boiteux, V.; Bach, C.; Sagres, V.; Hemard, J.; Colin, A.; Rosin, C.; Munoz, J. F.; Dauchy, X., Analysis of 29 per- and polyfluorinated compounds in water, sediment, soil and sludge by liquid chromatography-tandem mass spectrometry. *International Journal of Environmental Analytical Chemistry* **2016**, *96*, (8), 705-728.

75. Zhang, H.; Wen, B.; Hu, X.; Wu, Y.; Luo, L.; Chen, Z.; Zhang, S., Determination of fluorotelomer alcohols and their degradation products in biosolids-amended soils and plants using ultra-high performance liquid chromatography tandem mass spectrometry. *Journal of Chromatography A* **2015**, *1404*, 72-80.

Kaviarasan, S.; Ramamurthy, N.; Gunasekaran, P.; Varalakshmi, E.; Anuradha,
C. V., Induction of alcohol-metabolizing enzymes and heat shock protein expression by
ethanol and modulation by fenugreek seed polyphenols in Chang liver cells. *Toxicology Mechanisms and Methods* 2009, *19*, (2), 116-122.

77. Guan, Y.; Zhu, Q.; Huang, D.; Zhao, S.; Jan Lo, L.; Peng, J., An equation to estimate the difference between theoretically predicted and SDS PAGE-displayed molecular weights for an acidic peptide. *Scientific Reports* **2015**, *5*, (13370), 1-11.

APPENDIX



Fig. A1. PCR products of ssuD (1.2 kb), ssuE (0.6 kb), and tauD (0.9 kb).



Fig. A2. Agarose gel showing cut and uncut recombinant plasmids extracted from NEB 5- α competent *E. coli*.

Table A1. Expected plasmid DNA fragment sizes after restriction digestion. pET11a-ssuE was digested with XhoI (1 cut) and HpaI (2 cuts) or without enzymes (No cut). pET11a-ssuD and pET11a-tauD were digested with NotI-HF (1 cut) and HpaI (2 cuts) or without restriction enzymes (No cut).

Plasmid	1 cut (kb)	2 cuts (kb)								
pET11a-ssuE	6.3	1.8 and 4.5								
pET11a-ssuD	6.9	1.6 and 5.3								
pET11a-tauD	6.6	1.7 and 4.9								
5201>AGAGGA 5201>AGAGGA				-	-	-			-	
--	--	---	--	--	---	--	--	--	--	---
107>AGAGGA	TCGAGATCTCO TCGAGATCTCO TCGAGATCTCO	GATCCCGCGAA GATCCCGCGAA GATCCCGCGAA	ATTAATACGA(ATTAATACGA(ATTAATACGA(CTCACTATAGO CTCACTATAGO CTCACTATAGO	3GGAATTGTGI 3GGAATTGTGI 3GGAATTGTGI	AGCGGATAACA AGCGGATAACA AGCGGATAACA	ATTCCCCTCI ATTCCCCTCI ATTCCCCTCI	АGAAATAATT АGAAATAATT АGAAATAATT	TTGTTTAACT TTGTTTAACT TTGTTTAACT	<u>TT</u> >5300 TT>5300 TT>206
5301> AAGAAG 5301>AAGAAG 207>AAGAAG	GAGATATACA GAGATATACA GAGATATACA	FATGatgagca FATGatgagca FATGATGAGCA	ctgcattcgaa ctgcattcgaa CTGCATTCGAA	Aaccagaaca aaccagaaca AACCAGAACA	cgtcgctgac ccgtcgctgac ccgtcgctgac	ccctcgacaag ccctcgacaag CCCTCGACAAG	tteggeeeee tteggeeeee TTCGGCCCCC	acttcggcgc ACTTCGGCGC	cgagatcatc cgagatcatc CGAGATCATC	gg >5400 gg>5400 GG>306
5401>cotoos			-	, ,			+actacttct	-	+ +	aa>5500
5401>cctgga 307>CCTGGA	CGTCGCGTCC	gccaccgacga GCCACCGACGA	cgaggtggccg CGAGGTGGCCC	CGATCCGTT	cggcgttgaco CGGCGTTGACO	CGAACACAAGG CGAACACAAGG	tgetggttet TGCTGGTTC1	gcgcggtcag GCGCGGGTCAG	tcactgggcg TCACTGGGCG	ac>5500 AC>406
	*	*	*	*	*	*	*	*	*	*
5501>gcgagc	cacatcgagt	tcaatcaccat	ctcggccggct	gaccgcggg	gcatcccgtcg	cacqacaqcqq	cgacgtcgcg	caggaggtgt	acgcgctcga	ca >5600
5501>gcgagco 407>GCGAGCO	CACATCGAGT	tcggtcgccgt rCGGTCGCCGT	ctcggccggct cTCGGCCGGC	GACCGCGGGG	JCATCCCGTCC JCATCCCGTCC	cacgacagcgg CACGACAGCGG	CGACGTCGCC	CAGGAGGTGT.	acgcgctcga ACGCGCTCGA	ca>5600 CA>506
	*	*	*	*	*	*	*	*	*	*
5601>gccagg 5601>gccagg 507>GCCAGG	acaacggttto acaacggttto ACAACGGTTTO	cgccgacgtgt cgccgacgtgt CGCCGACGTGT	ggcacacgga ggcacacgga GGCACACGGA	cgtgacgttca cgtgacgttca CGTGACGTTCA	atgaagcggco atgaagcggco ATGAAGCGGCO	cgccgctgggg cgccgctgggg CGCCGCTGGGG	tcgatcctcc tcgatcctcc TCGATCCTCC	gtcccgtcgt gtcccgtcgt GTCCCGTCGT	cctcccgccg cctcccgccg CCTCCCGCCG	<u>ca</u> >5700 ca>5700 CA>606
	*	*	*	*	*	*	*	*	*	*
5701>cqqcqq	cgacaccaact	tgggccgacag	ccaactcgcgt	tacgagtcgc	tgtcgctgccg	ggtccggcaga	tgatcgatca	gctgaccgcg	gtgcacgacg	<u>gc</u> >5800
5701>cggcgg 607>CGGCGG	CGACACCAAC	Lgggccgacag IGGGCCGACAG	CCAACTCGCG	Lacgagtcgc TACGAGTCGC	tgtcgctgcco IGTCGCTGCCO	ggtccggcaga GGTCCGGCAGA	tgatcgatca TGATCGATCA	gctgaccgcg GCTGACCGCG	gtgcacgacg GTGCACGACG	gc>5800 GC>706
	*	*	*	*	*	*	*	*	*	*
5801>aaccgc	gagttcgggt	actacctcgcg	cagaagcgtg	aggcaaggg	aaacgtctgg	gacggtgagga	agtcaccgcg	ctcgtgcccg	tcgagcatcc	tg>5900
707>AACCGC	GAGTTCGGGT	ACTACCTCGCG	CAGAAGCGTG	GAGGCAAGGG	AAACGTCTGG	GACGGTGAGGA	AGTCACCGCG	CTCGTGCCCG	TCGAGCATCC	TG>806
	*	*	*	*	*	*	*	*	*	*
5901>tcgtgc	gggttcaccc					-				
		ggagacgggac	gcaagggaato	cttcgtgaac	ccgggcttcac	cgtcccacatc	gccggagtgt	ccgaggcgga	gagccgggga	at>6000
807>TCGTGC	gggttcaccco GGGTTCACCCO	ggagacggggac ggagacggggac GGAGACGGGAC	gcaagggaato gcaagggaato GCAAGGGAATO	cttcgtgaac cttcgtgaac CTTCGTGAAC	ccgggcttcac ccggggcttcac CCGGGCTTCAC	cgtcccacatc cgtcccacatc CGTCCCACATC	gccggagtgt gccggagtgt GCCGGAGTGT	ccgaggcgga ccgaggcgga CCGAGGCGGA	gagccgggga gagccggggga GAGCCGGGGA	<u>at</u> >6000 at>6000 AT>906
807>TCGTGC	gggttcaccco GGGTTCACCCO *	ggagacgggac ggagacgggac GGAGACGGGAC *	gcaagggaato gcaagggaato GCAAGGGAATO *	ttcgtgaaco ttcgtgaaco TTCGTGAACO	ccgggcttcac ccgggcttcac CCGGGCTTCAC *	cgtcccacatc cgtcccacatc CGTCCCACATC *	gccggagtgt gccggagtgt GCCGGAGTGT *	ccgaggcgga ccgaggcgga CCGAGGCGGA	gagccgggga gagccgggga GAGCCGGGGA *	<u>at</u> >6000 at>6000 AT>906 *
5901>tcgtgc 807>TCGTGC 5901>tcgtgc 5901>tcgtgc 277>TCGTGC	gggttcacco GGGTTCACCCC * gggttcaccc GGGTTCACCCC	ggagacgggac ggagacgggac GGAGACGGGGAC * ggagacgggac ggagacgggac GGAGACGGGAC	gcaagggaat gcaagggaat GCAAGGGAAT * gcaagggaat gcaagggaat GCAAGGGAAT	ttcgtgaac ttcgtgaac TTCGTGAAC * cttcgtgaac cttcgtgaac CTTCGTGAAC	ccgggcttcac ccgggcttcac CCGGGCTTCAC * ccgggcttcac ccgggcttcac CCGGGCTTCAC	cytcccacatc cgtcccacatc CGTCCCACATC * cgtcccacatc cgtcccacatc CGTCCCACATC	gccggagtgt gccggagtgt gccgGAGTGT * gccggagtgt gccggagtgt gccGGAGTG	ccgaggcgga ccgaggcgga cCGAGGCGGA * ccgaggcgga ccgaggcgga CCGAGGCGGA	gagccgggga gagccgggga GAGCCGGGGA * igagccgggga igagccgggga	at>6000 at>6000 AT>906 * aat>6000 at>6000 AT>376
5901>tcgtgc 807>TCGTGC 5901>tcgtgc 5901>tcgtgc 277>TCGTGC	gggttcaccc GGGTTCACCCC * gggttcaccc GGGTTCACCCC *	ggagacgggac ggagacgggac GGAGACGGGAC * ggagacgggac ggagacgggac GGAGACGGGAC *	gcaagggaat gcaagggaat GCAAGGGAAT * gcaagggaat gcaagggaat GCAAGGGAAT	ttcgtgaac ttcgtgaac TTCGTGAAC * <u>ttcgtgaac</u> ttcgtgaac CTTCGTGAAC	ccgggcttcac ccgggcttcac ccgggcttcac ccgggcttcac ccgggcttcac ccgggcttcac ccgggcttcac ccgggcttcac	x cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc	<pre>gccggagtgt gccggagtgt gccGGAGTGT * gccggagtgt gccggagtgt GCCGGAGTG? *</pre>	<pre>ccgaggcgga ccgaggcgga ccGAGGCGGA * ccgaggcgga ccGaggcgga fcCGAGGCGGA * * * * *</pre>	gagccgggga gagccgggga GAGCCGGGGA * gagccgggga GAGCCGGGGA *	at>6000 at>6000 AT>906 * at>6000 at>6000 AT>376 *
5901>tcgtgc 807>TCGTGC 5901>tcgtgc 5901>tcgtgc 277>TCGTGC 6001>cctcga 377>CCTCGA	gggttcaccc GGGTTCACCCC * gggttcaccc GGGTTCACCCC * <u>cttcctctac</u> cttcctctac CTTCCTCTACC	<pre>gagacgggac GGAGACGGGAC * ggagacgggac gGAGACGGGAC ggagacgggac GGAGACGGGAC * gcgcacctgac gcgcacctgac GCGCACCTGAC</pre>	gcaaggqaat gcaaggqaat gcaaggqaat gcaaggqaat gcaaggqaat gcaaggqaat gcaaggqaat gcaagcqagat caaacccqag caaacccqag cAACCCCAG	ttcgtgaac ttcgtgaac TTCGTGAAC * ttcgtgaac cttcgtgaac cttcgtgaac cttcgtgaac * cacatcgtgc cacatcgtgc cACATCGTGC	ccgggcttcar ccgggcttcar ccGGGCTTCAr * ccgggcttcar ccGGGCTTCAr * ggcaccgctgr ggcaccgctgr ggcaccgctgr	egteccacate cgtecccacate cgtec	<pre>gccgqaqtgt gccgqaqtgt GCCGGAGTGT * gccgqaqtgt gccgqaqtgt GCCGGAGTG? * acctcgtgtt ACCTCGTGT?</pre>	CCGAGGGGGGGG CCCAAGGCGGA * tcCGAGGCGGA tcCGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	gagccggqga gagccgggga GAGCCGGGGA * Mgagccgggga GAGCCGGGGA \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$	at>6000 at>6000 AT>906 * at>6000 at>6000 at>6000 at>6000 at>6100 ccg>6100 ccg>476
500)>tcgtgc. 807>TCGTGCC 590)>tcgtgc. 277>TCGTGCC 6001>cctcga 6001>cctcga 377>CCTCGA 6101>cactac. 6101>cactac.	gggttcaccc; gggttcaccc; gggttcaccc; gggttcaccc; cttcctcac; cttccttac; cttccttac; cttccttac; cttccctac; cttcc; cttcc; cttcc; cttcc; cttcc; cttcc; cttcc; cttc; cttc; cttcc; cttc; cttc; cttc; cttc; cttc; cttc; cttc; ctc; cttc; cttc; cttc; cttc; cttc; cttc; cttc; cttc; cttc; cttc; cttc; cttc; ctc; ctc; ctc; cttc;	<pre>gagaacqqqac ggagacqqqac ggagacqqqac ggagacqqqac gGAGACGGGAC * gcqcaccqqac gcgcacctqac gcgcacctqac GCGCACCTGAC * actacqqaacc actacqgaacc</pre>	gcaagggaat GCAAGGGAATG * gcaagggaat GCAAGGGAATG * caaacccgag CAAAGGGAATG * caacccgag CAAACCCGAG * cagcaccgcg cagcaccgcg	ttcgtgaac ttcgtgaac TTCGTGAAC * ttcgtgaac CTTCGTGAAC * cacatcgtgc cacatcgtgc cacatcgtgc cacatcgtgc cacatcgtgc tcatgcaccg tcatgcaccg	ccqqqcttca ccqqqcttca ccGGGGCTTCA * ccqqqcttca ccqqgcttca ccGGGCTTCA * gqcaccqctq ggcaccqctq GGCACCGCTGA * catcacqctq catcacqctq catcacqctq	cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccccacatc cgtcccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtccccccacatc cgtcccccccacatc cgtcccccccacatc cgtcccccccacatc cgtcccccccacatc cgtcccccccacatc cgtcccccccacatc cgtcccccccacatc cgtcccccccccacatc cgtcccccccccc	<pre>gccgqaqtqt gccgqaqtqt scccGGAGTGT * gccgqaqtgt gccgqaqtgt gccGGAGTGC * acctcgtgtt Acctcgtgtt AcctcGTGTC * gcccqtcgq gcccqtcgq gcccqtcgq gcccqtcgq gcccqtcgq</pre>	<pre>cccgaggcgga cccgaggcgga cccgaggcgga cccgaggcgga cccgaggcgga cccGAGGCGGA * tccgaggcgga cccGAGGCGGA * tgtgggacaac cGTGGGACAAC * tccgcagtagG cccgcagtagG cccgcagtagG cccgcagtagG</pre>	gagccqqqqa gagccgqqqa (GAGCCGGGGA * yqqqccqqqa (GAGCCGGGGA * * ccqcaqcaccq ccgcaqcaccq ccq ccgcaqcaccq ccq ccq ccq ccq ccq ccq ccq ccq	at>6000 at>6000 AT>906 * at>6000 AT>376 * * teg>6100 Ccg>6100 Ccg>6100 Ccg>6100 Ccg>6100 Ccg>6100 Ccg>6100 Ccg>6200 TCG>6200
500)>tcgtgc. 807>TCGTGCI 590)>tcgtgc. 590)>tcgtgc. 277>TCGTGCI 6001>cctcga. 377>CCTCGA 6101>cactac. 6101>cactac. 477>CACTAC	gggttcaccc; GGGTTCACCCC * gggttcaccc; GGGTTCACCCC; * cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccctcac; cttccccca; cttccccc; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccc; cttcc; cttcc; cttcc; cttcc; cttcc; cttcc; cttcc; cttcc; cttcc; cttcc; cttc; cttcc; cttc; ct; ct	<pre>gagacgqgac ggagacgqgac GGAGACGGGAC * gagacgqgac GGAGACGGGAC * gcgcacctgac gcgcacctgac GCGCACCTGAC * actacggaacc ACTACGGAACC</pre>	gcaagggaat GCAAGGGAATG * gcaagggaat GCAAGGGAATG * CAAGGGAATG * caaacccgag CAAACCCGAG * cagcaccgag CAAACCCGAG	ttcgtgaac ttcgtgaac TTCGTGAAC * ttcgtgaac CTTCGTGAAC * cacatcgtgc CACATCGTGC * tcatgcaccg tcatgcaccg	ccgggcttcad ccgggcttcad ccGGGCTTCAd cCGGGCTTCAd * ccgggcttcaa cCGGGCTTCAd * ggcaccgctg GGCACCGCTGd * catcacgctg catcacgctg cATCACGCTGd	cgtcccacatc cgtccccacatc cgtccccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtccccacatc cgtccccacatc cgtccccccc cgtccccccc cgccccccc cgcc cgcccccccc	<pre>gccgqaqtqt gccgqaqtqt cccGGAGTGT * gccgqaqtgt gccgqaqtgt GCCGGAGTG' * acctcqtqtt Acctcqtqtt Acctcqtqt acctcqtqt acctcqtqt acctcqtqt Acctcqt Acctcqtqt Acctcqt Acctqqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctcqt Acctqt Acctcqt ACCtqt ACCTC ACCTC ACCTC ACCTC ACCTC ACCTC AC</pre>	<pre>cccgagccga cccgagccga cccgagccga cccgagccga cccgagccga cccgagcgcga cccGAGCCGAA * tccgagcgcga cccGCAGCACAAC * cccgcagtagC cccgcagtagC cccGCAGTAGG </pre>	qaqccqqqaa gaqccqqqaa cAGCCGGGGA * uqaqccqqqaa iqaqccqqqaa iqaqccqqqaa cGAGCCGGGGA * ccqcaqcacqq ccGCAGCACGG cCTAGCATGAC cCTAGCATGAC	at>6000 at>6000 AT>906 * at>6000 AT>376 * tcg>6100 ;cg>6100 ;cg>6100 ;cg>6100 ;cg>6100 ;cg>6200 ;cg>6200 ;rg>6200
5901>tcgtgc. 807>TCGTGCI 5901>tcgtgc. 5901>tcgtgc. 277>TCGTGCI 6001>cctcga. 377>CCTCGA. 6101>cactac. 477>CACTAC.	gggttcaccc; GGGTTCACCCC * gggttcaccc; GGGTTCACCCC * cttcctctac; cttcctctac; cttcctctac; cttcctctac; cttcctctac; cttcctctac; cttcctcac; cttcctcac; cttcctcac; cttcccctac; cttcccctac; cttcccctac; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccca; cttccccc; cttccccc; cttccccc; cttccccc; cttccccc; ctcc; ctcccc; ctcc; ctcc; ctcc; ctcc; ctcc; ctcc; ctcc; ctcc; ctc; ctcc; ctc; ctcc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; ctc; cc; c	<pre>gagacgqgac ggagacgqgac ggagacqqgac ggagacqqgac ggagacqqgac ggagacqqgac gcgcacctgac gcgcacctgac gcgcacctgac gcgcacctgac actacggaacc actacggaacc actacggaacc actacggaacc</pre>	gcaagggaat GCAAGGGAATG * gcaagggaat gcaagggaat GCAAGGGAATG * caaacccgag CAACCGGAG * cagcaccgcg CAACCGCGC *	ttcgtgaac ttcgtgaac TTCGTGAAC * cttcgtgaac ttcgtgaac cttcgtgaac cttcgtgaac cttcgtgaac ctcgtgaac ctcgtgaac ctcgtgaac ttcatgcac ttcatgcac ttcatgaac ttca	ccgggcttcac ccgggcttcac ccGGGCTTCAC * ccgggcttcac ccGGGCTTCAC * ggcaccgctg GGCACCGCTGC * catcacgctg CATCACGCTGC *	cgtcccacatc cgtccccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtcccacatc cgtccccacatc cgtccccacatc cgtccccacatc cgtccccacatc cgtccccacatc cgtccccacatc cgcccccgcg cg cccccccccc	<pre>gccgqaqtqt gccgqaqtqt ccCGGAGTG7 * gccgqaqtqf gccgqaqtgf GCCGGAGTG7 * acctcqtgtt acctcqtgtt acctcqtgt * gcccqtcqtq gcccqtcqqq gcccqtcqqq gcccqtcqqq gcccqtcqqq gcccqtcqqq gcccqtcqqq gccqtcqqqq gccqtcqqqq gccqtcqqqq gccqtcqqqq gccqtqqqq gccqtqqqqqq gcqqqqqqq gcqqqqqqqq gcqqqqqqqqqq</pre>	<pre>ccgaqgcqqa cccgaqgcqga cccgaqgcqga tccgaqgcqga tccgaqgcqga tccgaqgcqga tccgaqgcqga tccGaGGCGGA * tgtgqacaac tGTGGGACAAC * tccgaqgagtagG tccGaCGACAAC * tccgaqgagtagG tccGACAAC * tccGaCGACAAC * tccGaCGACAAC * tccGACGACAAC * tccGACGACAAC * tcCGACGACAAC</pre>	gagccgggga gagccgggga cAGCCGGGA * igagccgggga GAGCCGGGGA * icgcagcaccg icgcagcagcag icgcagcagcagcagcag icgcagcagcagcag icgcagcagcagcag icgcagcagcagcagcag icgcagcagcagcagcag icgcagcagcagcagcagcag icgcagcagcagcagcagcag icgcagcagcagcagcagcagcagcag icgcagcagcagcagcagcagcagcagcagcagcagcagca	at>6000 at>6000 AT>906 * at>6000 lat\$6000 lat\$60000 lat\$60000 lat\$60000 lat\$60000 lat\$60000 l

Fig. A3-1. Alignment of sequencing result showing successful construction of recombinant plasmid pET11a-tauD. T7 promoter starts at 5230 bp. tauD insert starts at 5321 bp and ends at 6188 bp. Two primers were used in sequencing: pET11a-f (5'-ATA CCC ACG CCG AAA CAA GC -3') and tauD-f (5'-ACA GCC AGG ACA ACG GTT TC-3').

	*	*	*	*	*	*	*	*	*	*
5201>AGAGGAT	CGAGATCTC	GATCCCGCGAA	ATTAATACGA	CTCACTATAG	GGGAATTGTG	AGCGGATAACA	ATTCCCCTC	AGAAATAATT	TTGTTTAACT	TT >5300
5201>AGAGGAT	CGAGATCTC	GATCCCGCGAA	ATTAATACGA	CTCACTATAG	GGGAATTGTG	AGCGGATAACA	ATTCCCCTC	AGAAATAATT	TTGTTTAACT	TT>5300
109>AGAGGAT	CGAGATCTC	GATCCCGCGAA	ATTAATACGA	TCACTATAG	GGGAATTGTG	AGCGGATAACA	ATTCCCCTC	AGAAATAATT	TTGTTTAACT	TTT>208
1000 1101100111	0011011010									
		*	*	*		*	*	*	*	
E201>8808800			-	-	-	-		-	-	
5301>AAGAAGG	AGATATACA	TATGatgagca	tegactteta	etggegeate	gggatggaag	gtgatcacgcg	tcgttgcgca	redeededded	ctacaaccgo	2gg>5400
5301>AAGAAGG	AGATATACA	TATGatgagca	tegaetteta	ctggcgcatc	gggatggaag	gtgatcacgcg	tcgttgcgca	redeededded	ctacaaccgo	2gg>5400
209>AAGAAGG	AGATATACA	TATGATGAGCA	ICGACTICIA	CTGGCGCATC	GGGATGGAAG	GTGATCACGCG	TCGTTGCGCI	ACGCCGCGGCG	GCTACAACCGG	GG>308
	*	*	*	*	*	*	*	*	*	*
5401>cgacgcc	ggggggcac	cgacccggcaa	categeteeg	gcgatccggg	gcgggggaacto	cgacggctacg	gatacatcg	accacatggco	gcggtggcga	ag>5500
5401>cgacgcc	ggggggcac	cgacccggcaa	categeteeg	gcgatccggg	gcgggggaacto	cgacggctacg	gatacatega	accacatggco	gcggtggcga	ag>5500
309>CGACGCC	GGGGGGCAC	CGACCCGGCAA	CATCOCTCCO	GCGATCCGGG	GCGGGGGAACT	CGACGGCTACO	GATACATCO	ACCACATGGCC	GCGGTGGCGI	AG>408
	*	*	*	*	*	*	*	*	*	*
5501>acateca	aatcaacca	aatteetaaae	antetactac	ratectteen	ataacaaaa	accoctage	caccacaac	acactaacca	ccaaaacca	ca>5600
5501>gogtoog	agtoggoog	aattootaaaa	agtetactact	agtoottooo	gatgacggaa	raccostagge	caccogogggo	acactacco	Cogagacca	20225600
400>cccmccc	Acmedeece	commocrace			CAMCACCCAN		Checcecce		CCCACACCAC	
403/6061006	AGICGGCCG	GGIICCIGGGC	3610160160	GICCIICCO	GATGACGGAA	SACCCCIGGGC	CACCOCOOCI		CCGAGACCAC	CA-200
	*	*	*	*	*	*	* .	*	*	*
5601>cgtaccg	tttcatggt	cgccttccagc	ccgggttcct	gcacccggta	caggcggcca	ggatgtccgcg	agcctgcaa	cgcgccaccgg	aggccggct	igt >5700
5601>cgtaccg	tttcatggt	cgccttccagc	ccgggttcct	gcacccggta	caggcggcca	ggatgtccgcg	agcctgcaa	cgcgccaccgg	aggccggct	gt>5700
509>CGTACCG	TTTCATGGT	CGCCTTCCAGC	CCGGGTTCCT	GCACCCGGTA	CAGGCGGCCA	GGATGTCCGCC	AGCCTGCAA	GCGCCACCGG	AGGCCGGCTC	GT>608
	*	*	*	*	*	*	*	*	*	*
5701>ctacaac	atcatcago	aacaacaacaa	Laccotcaa	stataataaa	gagataaagto	cgatcacgaco	accactaca	ccacacatec	gagttected	ac>5800
5701>ctacaac	atcatcago	99099099099	tracatas	atataataaa	gagataaagt	agatoacqueg	accortacq	cogcacgtoc	gagttooto	120 5800
600>0000000000	AMCAMCACC		TCCCCCTCAN		CACAMAAAC	CONTRACTOR	ACCCCTTACC		Chammeemee	100 709
009×CIACAAC	AICAICAGO		IGCCCCICAA	101001000	GAGAIAAAGI	CGATCACGACC	ACCOUTACO	CCGCACGICC	GAGIICCICC	SAC>700
	*	*	*	*	*	*	*	*	*	*
5801>gtcctcc	gcggggtct	gggacggagag	ccattcgact	LCGaCGGCCG	gttcttccgca	accgacggggg	cgccctggc	cccggtctgg	ccggacaaco	:gt>5900
5801>gtcctcc	gcggggtct	gggacggagag	ccattcgact	Logaoggoog	gttcttccgca	accgacggggg	cgccctggc	cccggtctgg	ccggacaaco	gt>5900
709>GTCCTCC	GCGGGGGTCT	GGGACGGAGAG	CCATTCGACT	ICGACGGCCG	GTTCTTCCGCI	ACCGACGGGGG	CGCCCTGGC	CCCCGGTCTGG	CCGGACAACO	CGT>808
	*	*	*	*	*	*	*	*	*	*
5901>tccccga	* ggtgtactt	* ctccggctcgt	*	*	* acaaacaccca	* actcggactac	*	*	*	* ct>6000
5901> <u>tccccga</u>	* ggtgtactt	* ctccggctcgt	* ccgcggcagc	*	* gcgggcaccca	* actcggactac	*	*	*	* sct>6000
5901> <u>tccccga</u> 5901>tccccga	* ggtgtactto	* ctccggctcgt	* ccgcggcagc ccgcggcagc	* cgtcgccgcc cgtcgccgcc	* gcggggcaccca gcggggcaccca	* actcggactac actcggactac	* tacctgtcc	* :ggctcgagcc :ggctcgagcc	* gtacgcagad	* sct>6000 sct>6000
5901> <u>tccccga</u> 5901>tccccga 809>TCCCCGA	* ggtgtactto GGTGTACTTO	* ctccggctcgt ctccggctcgt CTCCGGCTCGT	* ccgcggcagco ccgcggcagco CCGCGGCAGC	* cgtcgccgcc cgtcgccgcc CGTCGCCGCC	* gcggggcaccca gcggggcaccca GCGGGCACCCA	* actcggactac actcggactac ACTCGGACTAC	* tacctgtcc TACCTGTCC	* ggctcgagcc ggctcgagcc GGCTCGAGCC	* gtacgcagac gtacgcagac GTACGCAGAC	* 2010 2010 2010 2010 2010 2010 2010 201
5901> <u>tccccga</u> 5901>tccccga 809>TCCCCGA	* ggtgtactto GGTGTACTTO *	* ctccggctcgt ctccggctcgt CTCCGGCTCGT *	* <u>ccgcggcagc</u> ccgcggcagc CCGCGGCAGC *	* cgtcgccgcc cgtcgccgcc cGTCGCCGCC *	* gcgggcaccca gcggggcaccca GCGGGCACCCA *	* actcggactac actcggactac ACTCGGACTAC *	* tacctgtcc tacctgtcc TACCTGTCC	* cggctcgagcc cggctcgagcc rGGCTCGAGCC *	* gtacgcagac gtacgcagac GTACGCAGAC *	* cct>6000 cct>6000 ccT>908 *
5901> <u>tccccga</u> 5901>tccccga 809>TCCCCGA	* ggtgtactto ggtgtactto GGTGTACTTO *	* ctccggctcgt ctccggctcgt cTCCGGCTCGT *	* ccgcggcagc ccgcggcagc CCGCGGCAGC *	* cgtcgccgcc cgtcgccgcc cGTCGCCGCCC *	* gcgggcaccca gcgggcaccca GCGGGCACCCA *	* actcggactac actcggactac ACTCGGACTAC *	* tacctgtcc tacctgtcc TACCTGTCC	* cggctcgagcc cggctcgagcc rGGCTCGAGCC *	* gtacgcagad gtacgcagad GTACGCAGAG *	* ct>6000 ct>6000 cT>908 *
5901>tccccga 5901>tccccga 809>TCCCCGA 5901>tccccga	* ggtgtactte ggtgtactte GGTGTACTTe * ggtgtactt	* <u>ctccggctcgt</u> ctccggctcgt cTCCGGCTCGT * <u>ctccggctcgt</u>	* ccgcggcagca ccgcggcagca cCGCGGCAGCA * ccgcggcagcagca ccgcggcagcagcagcagcagcagcagcagcagcagcagc	* cgtcgccgcc cgtcgccgcc cGTCGCCGCCC * cgtcgccgcc	* gcgggcaccca gcggggcaccca gCGGGCACCCJ * gcggggcacccc	* actcggactac actcggactac ACTCGGACTAC * actcggactac	* tacctgtco TACCTGTCC *	* cggctcgagcc cggctcgagcc GGCTCGAGCC * tggctcgagcc	* gtacgcagac GTACGCAGAC * gtacgcagac	* cct>6000 cct>6000 cct>908 * cct>6000
5901> <u>tccccga</u> 5901>tccccga 809>TCCCCGA 5901> <u>tccccga</u> 5901>tccccga	* ggtgtactt GGTGTACTT * ggtgtactt ggtgtactt	* ctccggctcgt ctccggctcgt cTCCGGCTCGT * ctccggctcgt ctccggctcgt	* ccgcggcagc ccGcGGCAGC * ccgcggcagc ccgcggcagc	* cgtcgccgcc cgtcgccgcc cGTCGCCGCCC * cgtcgccgcc cgtcgccgcc	* gcgggcaccca gcgggcaccca gcgggcaccca * gcgggcaccca gcgggcaccca	* actcggactac actcggactac ACTCGGACTAC * actcggactac	* tacctgtcc TACCTGTCC * <u>tacctgtcc</u> tacctgtcc	* cggctcgagcc cggctcgagcc cGGCTCGAGCC * tggctcgagcc	* gtacgcagac gtacgcagac GTACGCAGAC * cgtacgcaga gtacgcaga	* cct>6000 cct>6000 cct>908 * cct>6000 cct>6000 ccc>202
5901> <u>tccccga</u> 5901>tccccga 809>TCCCCGA 5901> <u>tccccga</u> 5901>tccccga 124>TCCCCGA	* ggtgtactt GGTGTACTT * ggtgtactt GGTGTACTT	* ctccggctcgt ctccggctcgt ctccGGCTCGT * ctccggctcgt ctccggctcgt cTCCGGCTCGT	* ccgcggcagc ccgcggcagc ccGCGGCAGC * ccgcgggcagc cCGCGGGCAGC	* cgtcgccgcc cgtcgccgcc cGTCGCCGCCC * cgtcgccgcc cGTCGCCGCC	* gcgggcaccc; gcgggcaccc; gcgggcaccc; * gcggggcaccc; gcgggcaccc; gcgggcaccc;	* actcggactac actcggactac ACTCGGACTAC * actcggactac ACTCGGACTAC	* tacctgtcc tacctgtcc TACCTGTCC * tacctgtcc TACCTGTCC	* ggctcgagcc regctcgagcc regctcgagcc * tggctcgagcc tggctcgagcc regctcgagcc	* gtacgcagac gtacgcagac GTACGCAGAC * gtacgcagac gtacgcagac CGTACGCAGAC	* <u>set</u> >6000 cct>908 * <u>sect</u> >6000 cct>6000 cct>223
5901> <u>tccccga</u> 5901>tccccga 809>TCCCCGA 5901> <u>tccccga</u> 5901>tccccga 124>TCCCCGA	* ggtgtactt GGTGTACTT * ggtgtactt GGTGTACTT	* ctccggctcgtc ctccggctcgtc ctccggctcgtc ctccggctcgt ctccggctcgt ctccggctcgt	* ccgcggcagc ccgcggcagc ccGCGGCAGC * <u>ccgcggcagc</u> ccgcggcagc CCGCGGCAGC	* cgtcgccgcc CGTCGCCGCCC * cgtcgccgcc cgtcgccgcc CGTCGCCGCC	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcacccc gcgggcacccc GCGGGCACCC	* actoggactac ACTCGGACTAC * actoggactac ACTCGGACTAC	* tacctgtcc TACCTGTCC * tacctgtcc tacctgtcc TACCTGTCC	* cggctcgagcc rGGCTCGAGCC * tggctcgagcc rGGCTCGAGCC rGGCTCGAGCC	* gtacgcagac GTACGCAGAC * cgtacgcagac cgtacgcagac CGTACGCAGAC	* 221>6000 2CT>908 * 221>6000 2CT>223
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 124>TCCCCGA	* ggtgtactt GGTGTACTT * ggtgtactt GGTGTACTT	* ctccggctcgt ctccGGCTCGT * ctccggctcgt ctccggctcgt ctccGGCTCGT	* ccgcggcagc ccGCGGCAGCA * ccgcggcagc ccgcggcagc cCGCGGCAGC	* cgtcgccgcc cGTCGCCGCCG * cgtcgccgcc cgtcgccgcc cGTCGCCGCC	* gcgggcaccci gcgggcaccci gcGGGCACCCI * gcgggcaccci gcgggcaccci gcGGGCACCC	* actcggactac ACTCGGACTAC * actcggactac actcggactac ACTCGGACTAC	* tacctgtcc TACCTGTCC * tacctgtcc tacctgtcc TACCTGTCC	* cggctcgagcc rGGCTCGAGCC * tggctcgagcc tggctcgagcc tggctcgagcc rGGCTCGAGCC	* gtacgcagad GTACGCAGAG * cgtacgcagad cgtacgcagad CGTACGCAGAG	* ect>6000 cct>908 * ect>6000 cct>6000 cct>223
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 124>TCCCCGA	* ggtgtactt GGTGTACTT * ggtgtactt GGTGTACTT *	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt *	* ccgcggcagc ccgcggcagc ccgcggcagc ccgcggcagc ccgcggcagc ccgcggcagc ccgcggcagc ccgcggcagc *	* cgtcgccgcc cgtcgccgcc * cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc *	* gcgggcaccc; gcGGGGCACCC; * gcgggcaccc; gcGGGCACCC; *	* actcggactac actcggactac ACTCGGACTAC * actcggactac ACTCGGACTAC *	* tacctgtcc TACCTGTCC * tacctgtcc tacctgtcc TACCTGTCC	* ggctcgagcc regctcgagcc * tggctcgagcc * tggctcgagcc tggctcgagcc * *	* gtacgcagac gtacgcagac GTACGCAGAC * gtacgcagac GTACGCAGAC *	* cct>6000 cct>6000 cct>908 * cct>6000 cct>6000 cct>223 *
5901>tccccqa 5901>tccccqa 809>TCCCCGA 5901>tccccqa 5901>tccccqa 124>TCCCCGA	* ggtgtactt ggtgtactt ggtgtactt	* ctccggctcgt ctccggctcgt * ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga	* ccgcgggcagc ccGcGGCAGC * ccgcgggcagc ccGcGGCAGC * gcagtccgcg	* cgtcgccgcc cGTCGCCGCC * cgtcgccgcc cgtcgccgcc cGTCGCCGCC * aaactcggggc	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcggcaccci gcggcaccci	* actcggactac actcggactac * actcggactac actcggactac actcggactac actcggactac actcggactac gttcgcggtgg	* tacctgtcc/ TACCTGTCC * tacctgtcc tacctgtcc TACCTGTCC *	* ggctcgagcc rGGCTCGAGCC * tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc	* gtacgcagac GTACGCAGAC * ggtacgcagaa ogtacgcagaa GTACGCAGAC * * ggtacgcagaa	* cct>6000 cct>6000 cct>908 * cct>6000 cct>223 * ggcc>61000
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA	* ggtgtactt: ggtgtactt: ggtgtactt: ggtgtactt: ggtgtactt: ggtgtactt: aaattcgac aaattcgac	* atcaggetegt atcaggetegt creceggetegt atcaggetegt atcaggetegt creceggetegt aggegtacgaga ggegtacgaga	* gcagtccgcg ccgcggcagcc ccgcggcagcc ccgcggcagcc ccgcggcagcc ccgcggcagcc ccgcggcagcc *	* cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc	* acagagcaccci acci acagagcaccci acci acagagcaccci acci acci acagagcaccci acci acci acagagcaccci acci	* actcggactac actcggactac actcggactac actcggactac ACTCCGACTAC * actcggactac actcggactac actcggactac	* ttacctgtcc TACCTGTCC * tacctgtcc ttacctgtcc TACCTGTCC * ygqatcgaca	* agotogagoo agotogagoo recorceAcco * tagotogagoo recorceAcco * tagotogagoo recorceAcco * tagotogagoo * * * * * * * * * * * * *	* gtacgcagac gtacgcagac GTACGCAGAC * gtacgcagac GTACGCAGAC * contactor * contactor * contactor * contactor * contactor * * contactor * * contactor * * * * * * * * * * * * *	* set>6000 cct>908 * set>6000 cct>6000 cct>223 * gac>6100 gac>6100
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> qcqqaaa 6001> qcqqaaa 224>CCCGAAA	* ggtgtactt: GGTGTACTT * ggtgtactt: GGTGTACTT * agatgtactt: GGTGTACTT * aaattcgac AAATTCGAC	* ctccggctcgt ctccgGCTCGT * ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga gGCGTACGAGA	* ccgcggcagca ccgcggcagc ccgccggcagc ccgcggcagc ccgcggcagc ccgccgcgagcagc ccgcggcagc ccgccgcgcagc gcagtccqcg gcagtccqcg gcagtccqcg	* cgtcgccgcc CGTCGCCCCCC * cgtcgccgcc cGTCGCCGCC cGTCGCCGCC cGTCGCCGCC * aaactcgggc AAACTCGGGC	* gcgggcaccc: gcgggcaccc: gcgggcaccc: gcgggcaccc: gcggcacccca gccacccccaa gcacccccaa gcacccccaa	* actcggactac AcTCGGACTAC * actcggactac ACTCGGACTAC * gtcgggactac gtcgcggtgg GTTCGCGGTG	* ttacctgtcci TACCTGTCC' * ttacctgtcc ttacctgtcc TTACCTGTCC' * ggatcgaca cGGATCGACA	* cggctcgagcc cggctcgagcc cgGctcGAGCC * tggctcgagcc rGGCTCGAGCC * tggctcgagcc tggctcgagcc rGGCTCGAGCC CGTCGCCCCG	* sgtacgcagac GGTACGCAGAC * sgtacgcagac GGTACGCAGAC * sgtacgcagac GGTACGCAGAC * sgtacgcagac CGCACCGGAAC	* cct>6000 ccT>908 * cct>6000 cct>6000 cct>223 * gac>6100 gac>6100 gac>323
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> <u>gcqqaaa</u> 6001>gcqgaaa 224>GCGGAAA	* ggtgtactt: ggtgtactt: ggtgtactt: ggtgtactt: ggtgtactt: acaattcgac AAATTCGAC	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctgt ctccggctgt ggcgtacgaga ggcgtacgaga ggcgtacgaga	* ccgcggcagca cccccGcGcAcCc * ccgcggcagcagc ccgcggcagc ccgccGcCACCC * gcagtccgcg gcagtccgcg gcagtccgcg gcAgtccgcg	* cgtcgccgcc cGTCGCCGCCCC * cgtcgccgcc cgtcgccgcc cGTCGCCGCCCCC * aaactcgggc AAACTCGGGC	* gcgggcaccc: gcgggcaccc: gcgggcaccc: gcgggcaccc: gcgggcaccc: gcggcaccccaa: gcacccccaa: gcacccccaa:	* actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac gttcgcggtgg gttcgcggtgg	* ttacctgtcc' tracctgtcc' * ttacctgtcc ttacctgtcc ttacctgtcc ttacctgtcc * tggatcgaca cggatcgaca cggatcgaca	* cggctcgagcc cggctcgagcc cggctcgagcc * tggctcgagcc tggcc	* gtacgcagac GTACGCAGAC * gtacgcagac GTACGCAGAC * coacaccgaac CCACACCGAAC	* cct>6000 ccT>908 * cct>6000 cct>6000 ccT>223 * gac>6100 gac>6100 gac>323
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> <u>gcqqaaa</u> 224>GCGGAAA	* ggtgtactt GGTGTACTT * ggtgtactt GGTGTACTT * agtgtactt GGTGTACTT * * aaattcgac AAATTCGAC	* ctccggctcgt ctccggctcgt * ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga ggcgtacgaga	* ccgcggcagca ccgcggcagc ccccccGCGCAGCC * ccgcggcagcagc cccCCCGCGCAGCC * gcagtccgcg gcagtccgcg GCAGTCCGCG	* cgtcgccgcc cGTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCC cGTCGCCCCCCC * aaactcgggc AAACTCGGGC	* gcgggcaccc * gcgggcaccc gcGGGCACCCC * gcgggcaccc GCGGGCACCCC * gcacccccaa gcacccccaa GCACCCCCAA	* actcggactac actcggactac * actcggactac Actcggactac Actcggactac Actcggactac gttcgcgqtgc gttcgcgqtgc	* ttacctgtcci ttacctgtcci * ttacctgtcc ttacctgtcc ttacctgtcc * ttacctgtcc * ttacctgtcc * ttacctgtcc * ttacctgtcc * ttacctgtcc * * ttacctgtcc * * ttacctgtcc * * ttacctgtcc * * * * * * * * * * * * *	* cggctcgagcc cggctcgagcc cgcctcgagcc * tggctcgagcc cggctcgagcc tggctcgagcc tggctcgagcc * tggctcgagcc cgccccccccc *	* sgtacgcagac GGTACGCAGAC * sgtacgcagac GGTACGCAGAC * scoracgcagac CCACACCGAAC	* cct>6000 ccT>908 * cct>6000 cct>6000 ccT>223 * gac>6100 gac>6100 GAC>323
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>gcqgaaa</u> 6001>gcqgaaa 224>GCGGAAA	* ggtgtactt: GGTGTACTTO * ggtgtactt: GGTGTACTTO * * maaattcgac AAATTCGAC *	* ctccggctcgt ctccggctcgt * ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga ggcgtacgaga s	* ccgcggcagcagc ccccccGcGcAcCc * ccgcggcagcagc ccgcggcagc ccgccgcGcAcCc * gcagtccgcg gcagtccgcg gcAGTCCCCCG *	* cgtcgccgcc cGTCGCCCCCCC * cgtcgccgcc cgtcgccgcc cGTCGCCGCCCCC * aaactcgggc AAACTCCGGCC *	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcggcaccccaa gcacccccaa gcacccccaa gcacccccaa *	* actcggactac ACTCGGACTAC * actcggactac ACTCGGACTAC * gttcgggatgc gttcgcggtgc GTTCGCGGGTGC *	* ttacctgtcci tracctgtcci * ttacctgtcc ttacctgtcc ttacctgtcc * tggatcgaca cGATCGACA *	* cggctcgagcc cggctcgagcc cggctcgagcc * tggctcgagcc tggct tggctggc tggct tggct tggct tggct tggc tg	* sgtacgcagac cGTACGCAGAC * sgtacgcagac cGTACGCAGAC * sgtacgcagac cGTACGCAGAC * cacaccgaac cCACACCGAAC *	* 201 201 201 201 201 201 201 201
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>gcqqaaa</u> 224>GCCGAAA 6101>gcqqcat	* ggtgtactt: GGTGTACTT * ggtgtactt: GGTGTACTT * aggtgtactt: GGTGTACTT * aaattcgac AAATTCGAC *	* ctccggctcgt ctccggctcgt * ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga ggcgtacgaga ggcgtacgaga ggcgtacgagag ggcgtacgag ggcg ggcg ggag ggcg ggag ggcg ggag ggcg ggag ggcg ggcg ggag ggcg ggag ggcg ggag ggag ggcg gga	* ccgcggcagca ccgcggcagca ccgcggcagca ccgcgggcagc ccgcgggcagc ccgccgcgcagca gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg	* cgtcgccgcc cGTCGCCCCCCC * cgtcgccgccgcc cGTCGCCCCCCC * aaactcgggc AAACTCCGGCC * tcgacaagga	* gcgggcaccc: * gcgggcacccc gcGgGcAcCCC GCGGGCACCCC * gcacccccaa gcAccccccaa GCACCCCCAA	* actcggactac actcggactac * actcggactac AcTCGGACTAC * actcggactac ACTCGGACTAC * attcgcggtgg GTTCGCGGTGC *	* ttacctgtcci ttacctgtcci * ttacctgtcc ttacctgtcc ttacctgtcc * ggatcgaca CGATCGACA *	* cggctcgagcc cggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc * tcgtcgcccga rcGCTCGCCCCC *	* sgtacgcagad ig	* 201 201 201 201 201 201 201 201
5901>tccccqa 5901>tccccqa 809>TCCCCGA 5901>tccccqa 5901>tccccqa 124>TCCCCGA 6001>qcqqaaa 224>CCCGGAA 6101>qcqqcat 6101>qcqqct	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * aaattcgac AAATTCGAC * gggcggagaa gggcggagaa	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga ggcgtacgaga ggcgtacgaga tccgagaagggc	* ccgcggcagca ccgcggcagc ccgccggcagc ccgccggcagc ccgccggcagc ccgccggcagc ccgccggcagc ccgccggcagc ccgccgcagc gcagtccgcg gcagtccgcg gcagtccgcg tgggccttcg	* cgtcgccgcc CGTCGCCCCCCC * cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc * cgtcgccgcgc * cgtcgccgcgc * tcgaccagga tcgaccagga	* gcgggcaccc: gcgggcaccc: gcgggcaccc: gcgggcaccc: gcgggcaccc. gcgcaccccaa: gcacccccaa: gcacccccaa: gcacccccaa: tgccgcacacc:	* actcggactac actcggactac ACTCGGACTAC * actcggactac ACTCGGACTAC * actcggactac ACTCGGACTAC * gttcgcggtg gtcgcggtg GTTCGCGGTGG * cggqcgqcgcg	* ttacctgtcci TTACCTGTCC' * ttacctgtcc ttacctgtcc ttacctgtcc ctacctgtcc ctacctgtcc ctacctgtcc ctacctgtcc ctacctgtcc ctacctgtcc * * ctacctgtcc * ctacctgtcc * * ctacctgtcc * * ctacctgtcc * * ctacctgtcc * * ctacctgtcc * * ctacctgtcc * * * ctacctgtcc * * * ctacctgtcc * * * ctacctgtcc * * * * * * * * * * * * *	* cggctcgagcc cggctcgagcc cGGCTCGAGCC * tggctcgagcc tggctcgagcc * tggctcgagcc cGCTCGACCCG * cgtcgggggggggggggggggggggggggggggggggg	* cgtacgcagac cGTACGCAGAC * cgtacgcagac cGTACGCAGAC * cgtacgcagac cGTACGCAGAC * ccacaccgaac ccACACCGAAC * ccacaccgaac ccACACCGAAC	* cct>6000 ccT>908 * cct>6000 ccT>908 * cct>6000 cct>6000 cct>6000 ccT>223 * gac>6100 gac>6100 gac>6100 gac>6200 *
5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 809>TCCCCGA 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>gcqgaaa</u> 224>GCCGAAA 6101> <u>gcqgcat</u> 6101> <u>gcqgcat</u> 6101> <u>gcqgcat</u>	* ggtgtactti ggtgtactti ggtgtactti ggtgtactti	* ctccggctcgt cTCCGGCTCGT * ctccggctcgt ctccggctcgt ctccGGCTCGT * ggcgtacgaga gGCGTACGAGA * tcgagaagggc tcgagaagggc	* ccgcggcagc: ccgcggcagc. * ccgcggcagcagc ccgcgggcagc ccgccgcGCAGC * * gcagtccgcg gcagtccgcg gcAGTCCGCG * tgggccttcg tgggccttcg	* cgtcgccgcc ccTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCCC * aaactcgggc AAACTCGGCC * tcgacaagga tcgacaagga tcgacaagga	* gcgggcaccci cocceccaccci gcgggcacccc gcgggcacccc gcGggcacccc gcGcGcCACCCC * gcacccccaa gcAccccccaa gcAccccccaa gcAccccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa	* actcggactac actcggactac AcTCGGACTAC * actcggactac actcggactac actcggactac gttcgcggtg GTTCGCGGTGC * cgggcggcgcg cgggcggcgcg	* ttacctgtcci ttacctgtcci ttacctgtcci ttacctgtcci ttacctgtcci * rgggtcggaca rgggtcgaca rgggtcgaca rgggcgattci rggggattci rggggattci rggggattci rggggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci rgggattci	* :ggctcgagcc :ggctcgagcc :ggctcgagcc :ggctcgagcc :tggctcgagcc :tggctcgagcc : : : : : : : : : : : : :	* igtacgcagad igtacgcagad igtacgcagad igtacgcagad igtacgcagad igtacgcagad igtacgcagad igtacgcagad igccgcatcad igcccgcatcad	* 2012 2017 20
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> <u>qcqqaaa</u> 6001> <u>qcqqaaa</u> 6001> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u>	* ggtgtactt: GGTGTACTT * ggtgtactt: GGTGTACTT * agatgtactt: GGTGTACTT * aaattcgac AAATTCGAC * gggggggaga GGGCGGAGA	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga gGcCTACGAGA * tcgagaagggc tcgagaagggc tcgagaagggc	* ccgcggcagca ccgcggcagc * ccgcggcagcagc ccgcggcagcagc ccgcgggcagc gcagtccgcg gcagtcgcg gcagtcg gcagtcgcg gcagtcgcg gcagtcgcg gcagtcgcg gcagtcgcg gcagtcgcg gcagtcgcg gcagtcgcg gcagtcg gcagtcgcg gcagtcgcg gcagtcgcg gcagtcgcg gcagtcg gcagtcgcg gcagtcgcg gcagtcg gcagtcg gcagtcg gcagtcgcg gcagtcg gcag gcagtcg	* cgtcgccgcc CGTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	* gcgggcacccc gcgggcacccc gcgggcacccc gcgggcacccc gcgggcacccc gcGcGcCcCcAa gcacccccaa gcaccccccaa gcacccccaa gcacccccaa gcacccccccaa gcaccccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccccaa gcacccccccaa gcaccccccccaa gcacccccccccc	* actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac ggtcgggggggggg	* ttacctgtcci TACCTGTCC' * ttacctgtcc ttacctgtcc TTACCTGTCC' * gggatcgaca CGATCGACA * gggggggattc ggggggattc GGGCGATTC	* cggctcgagcc cggctcgagcc cgcctcGAGCC * tggctcgagcc cggctcgagcc tggctcgagcc cgccccGAGCC * tggctcgagcc cgtcgccccgc cgtcggggggg cgtcggggggg cgtcggggggg	* sgtacgcagac GTACGCAGAG * sgtacgcagac CGTACGCAGAG * sgtacgcagac CGTACGCAGAG * sgcacgcagaa sgcccgcatcac SGCCGCATCA	* cct>6000 ccT>908 * cct>6000 cct>6000 cct>6000 cct>6000 cct>6000 cct>6100 gac>6100 gac>6100 gac>6200 GAC>323 *
5901>tccccqa 5901>tccccqa 809>TCCCCGA 5901>tccccqa 5901>tccccqa 124>TCCCCGA 6001>gcqgaaa 6001>gcqgaaa 224>GCCGGAAA 6101>gcqgcat 6101>gcqgct 6101>gcqgct	* ggtgtactt: GGTGTACTTV * ggtgtactt: GGTGTACTTV * aaattcgac .aaattcgac .AAATTCGAC * gggcggaga GGGCCGACA	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggggtacgaga gggg tcgagaaggggc tcgagaaggggc tcgagaaggggc tcgagaaggggc tcgagaaggggc tcgagaaggggc tcgagaaggggc tcgagaaggggc tcgagaaggggc tcgagaaggggc tcgagagggg tcgagaggg tcgagaggg tcgagaggg tcgagagggg tcgagagggg tcgagaggg tcgagaggg tcgagaggg tcgagagggg tcgagaggg tcgagaggg tcgagaggg tcgagaggg tcgagagggg tcgagagggg tcgagaggg tcgagagggg tcgagaggg tcgagagggg tcgagagggg tcgagaggg tcgagagggg tcgagagggg tcgagagggg tcgagagggg tcgagagggg tcgagggg tcgagggg tcgagagggg tcgagggg tcgagggg tcgagggg tcgagagggg tcgaggggg tcgagggggg tcgagggg tcgagggggg tcgaggggggg tc	* ccgcggcagcagc ccGccGcGcAGCC * ccgcggcagcagc ccgcggcagc ccGccGGcAGCC * gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg tgggccttcg TGGGCCTTCG	* cgtcgccgcc CGTCGCCCCCCC * cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc * aaactcgggc AAACTCGGCGC * tcgacaagga tcgacaagga TCGACAAGGA	* gcgggcaccc: gcGGGCACCCI * gcgggcaccc gcgggcaccc gcgggcaccc gcGGGCACCCCA * gcacccccaa gcacccccaa gcACCCCCAA * tgccgcaaac tgccgcaaac	* actcggactac actcggacgacgac actcggacgacgacgacgacgacgacgacgacgacgacgacga	* ttacctgtcci TTACCTGTCC: * ttacctgtcc ttacctgtcc ttacctgtcc * tgggatcgaca cGGATCGACA * * * * * * * * * * * * *	* cggctcgagcc cggctcgagcc cGCCTCGAGCC * tggctcgagcc tggctcgagcc tggctcgagcc cgtcgccggc cgtcggcgggggg cgtcggggggg cGTCGGGGCCG	* cgtacgcagac cGTACGCAGAC * cgtacgcagac cgtacgcagac cGTACGCAGAC * cgtacgcagac ccaccgaac ccaccgaac ccacccgaac ccacccgaac ccacccgaac ccacccgaac ccacccgaac ccaccacac ccaccacac ccaccacac ccaccacac ccaccacac ccaccacac ccaccacac ccaccacacac ccaccacacac ccaccacacac ccaccacacaca	* 2012 20
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>gcqqaaa</u> 224>GCGGAAA 6101> <u>gcqqcat</u> 6101> <u>gcqqct</u> 6101> <u>gcqqct</u>	* ggtgtactt: GGTGTACTT * ggtgtactt: GGTGTACTT * agtgtactt: GGTGTACTT * aaattcgac AAATTCGAC * gggcggaga GGGCGGACA *	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga GCGTACGACA * tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc	* ccgcggcagca ccgcggcagca ccgccggcagca ccgcggcagca ccgccgcGcAGCC * gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg tgggcattcg tgggccttcg tgggccttcg *	* cgtcgccgcc cGTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCCC * aaactcgggc AAACTCGGCC * tcgacaagga tcgacaagga tcgacaagga	* gcgggcaccci ccccccaccci gcgggcacccc gcgggcacccc gcGcGcCcCcCi * gcaccccccaa gcaccccccaa gcacccccaa gcacccccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccaa gcacccccaa gcacccccaa gcaccccccaa gcacccccccaa gcaccccccaa gcaccccccaa gcacccccaa gcaccccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcaccccaa gcacccccaa gcaccccaa gcaccccaa gcacccccaa gcaccccccaa gcacccccaa gcacccccaa gcaccccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcaccccaa gcacccccaa gcacccccaa gcaccacccaa gcacccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccaa gcaccaccac	* actcggactac actcggacgac actggacgacgac actggacgacgacgacgacgacgacgacgacgacgacgacgac	* ttacctgtcci ttacctgtcci * ttacctgtcc ttacctgtcc ttacctgtcc * tggatcgaca CGGATCGACA * Agggcgattc Agggcgattc Agggcgattc Agggcgattc Agggcgattc Agggcgattc Agggcgattc	* :ggctcgagcc :ggctcgagcc :ggctcgagcc :tggctcgagcc :ggctcgagcc : : : : : : : : : : : : :	* sgtacgcagac ig	* 201 201 201 201 201 201 201 201
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>qcqqaaa</u> 224>CCCGAA 6101> <u>qcqqaaa</u> 224>GCGGAA 6101> <u>qcqqct</u> 6101> <u>qcqqct</u> 324>GCGCCT	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * aaattcgac AAATTCGAC * gggcggaga GGGCGGAGA *	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga ggcgtacgaga ggcgtacgaga tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc	* ccgcggcagca ccgcggcagca * ccgcggcagc ccgccggcagc ccgccggcagc ccgccggcagc ccgccggcagc ccgccggcagc gcagtccgcg gcagtccgcg gcagtccgcg tgggcattcg tgggccttcg tgggccttcg	* cgtcgccgcc cGTCGCCCCCCC * cgtcgccgcc cgtcgccgcc cGTCGCCGCCCC * aaactcgggc aaactcgggc AAACTCGGCC * tcgacaagga TCGACAAGGA *	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcacccc gcggcaccccaa gcacccccaa gcacccccaa gcacccccaa Tgccgcaaacc tgccgcaaacc tgccgcaaacc tgccgcaaacc	* actcggactac actcggactac AcTCGGACTAC * actcggactac ACTCGGACTAC * actcggactac ACTCGGACTAC * actcggactac GGTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	* ttacctgtcci TTACCTGTCC' * ttacctgtcc ttacctgtcc TTACCTGTCC' * gggatcgaca cGGATCGACA * agggcgattc Agggcgattc Agggcgattc	* cggctcgagcc cggctcgagcc cgccrcGAGCC * tggctcgagcc tggctcgagcc * tcgtcgccggc tcgtcgccggc cgccCCCC * cgtcgggggg cgrcGGGGCCGG	* stagcagaa GTACGCAGAG * stagcagaa GTACGCAGAG * stagcagaa cacaccgaa cacaccgaa cacaccgaa cacaccgaa stagcagatca	* cct>6000 ccT>908 * cct>6000 ccT>223 * gac>6100 gac>6100 gac>6100 gac>6200 agg>6200 agg>6200 agg>6200 agg>6300
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>gcqqaas</u> 6001> <u>gcqqaas</u> 224>GCCGAAA 6101> <u>gcqqcat</u> 6101> <u>gcqqcat</u> 6101> <u>gcqqcat</u> 6101> <u>gcqqcat</u> 6201> <u>gctqqt</u>	* ggtgtactt: GGTGTACTT * ggtgtactt: GGTGTACTT * aaattcgac AAATTCGAC * gggcggaga GGCCCGACAA * acccgaacaa	* ctccggctcgt cTcCGGCTCGT * ctccggctcgt ctcCGGCTCGT * ggcgtacgaga ggcgtacgaga GGCGTACGAGA * tcgagaagggc tcgagaagggc ccAcAAAGGGC * catcaccggct	* ccgcggcagc ccGCGGCGCAGC * ccgcgggcagc gcagcagcagc gcagtccgcg gcagtccgcg gcagtccgcg tgggcettcg tgggcettcg tgggcettcg acggggacct * acgggacct * acggacct * acgggacct * acggacct	* cgtcgccgcc ccrcccccccc * cgtcgccgcc ccrcccccccc * aaactcgggc AAACTCCGCCC * tcgacaagga tcgacaagga tcgacaagga tcgacaagga ccacccccccccccccccccccccccccc	* gcgggcaccci ccccccaccci gcgggcacccc gcgggcacccc gccgccaccccac gcccccccaa gcacccccccaa gcaccccccaa gcaccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccc	* actcggactac actcggacgacgac actcggacgacgacgacgacgacgacgacgacgacgacgacga	* ttacctgtcci ttacctgtcci ttacctgtcc ttacctgtcc ttacctgtcc * gggtcgaca GGATCGACA * aggccgattc AGGCCGATTC *	* cgctcgagcc cggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc cgcccccccc * cgtcggcgcccgc cgcccccccc * cgtcggggggg cgrccGcccccc * cgccccccccccc * cgtcgggggggg cgrccGccccccc *	* gtacgcagad gtacgcagad gtacgcagad ggtacgcagad ggtacgcagad ggtacgcagad cGTACGCAGAG * cacacccgaad cCACACCGAAG * gcccgcatcad g	* 2012 2013 2015 2015 2015 2015 2015 2015 2015 2015 2015 2015 2015 201
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> <u>qcqqaaa</u> 224>CCCGAAA 6101> <u>gcqgacat</u> 6101> <u>gcqgcct</u> 6101> <u>gcqgcct</u> 6201> <u>gctqqt</u> 6201> <u>gctqqt</u>	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * aaattcgac AAATTCGAC * ggggcggaga GGGCGGACA * accccgaaca accccgaaca	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga gGcCTACGAGA * cgagagagggc tcgagaagggc tcgagaagggc tcgagaagggc ctcgacgagagggc ctcgacgagagggc ctcgacgagagggc ctcacggct catcaccggct catcaccggct	* ccgcggcagca ccgcggcagca * ccgcggcagcagc ccgcggcagcagc ccgcgggcagcagc ccgcgggcagcagc ccgcgggcagcagc gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg tgggccttcg * accggggacct * accggggacct	* cgtcgccgcc cGTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgcccccaa gcaccccccaa gcaccccccaa gcaccccc	* actcggactac actcgactac act	* ttacctgtcci TTACCTGTCC' * ttacctgtcc TTACCTGTCC' * ggatcgaca Ggatcgaca GGGATCGACA * agggcgattc GGGCGATTC * aggctgatc	* cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctggggcgg cgfcgggggggg cgfcgggggggg cgfcgggggggg	* cgtacgcagac cGTACGCAGAC * cgtacgcagac cGTACGCAGAC * cgtacgcagac cGTACGCAGAC * cacaccgaac cCACACCGAAC * cacccgacatcag ccccgcatcag ccccgcatcag ccccgcatcag ccccgcatcag ccccgcatcag ccccgcatcag ccccgcatcag cccccacaccgaa	* 2015 2015 2015 2015 2015 2015 2015 2015
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> <u>gcqgaaa</u> 224>GCCGAAA 6101> <u>gcqgcat</u> 6101> <u>gcqgcat</u> 6101> <u>gcqgcat</u> 6101> <u>gcqgcat</u> 6201> <u>gctqqt</u> 6201> <u>gctqqt</u> 6201> <u>gctqqt</u>	* ggtgtactti ggtgtactti ggtgtactti ggtgtactti ggtgtactti aaattcgac AAATTCGAC * gggcggaga gggcggaga gggcggaga acccgaaca ACCCGAACA	* ctccggctcgt cTCCGGCTCGT * ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga gGCGTACGAGA * tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc tcatcaccggct catcaccggct	* ccgcggcagc: ccgcggcagc ccgcggcagc ccgcggcagc ccgcggcagc ccgcggcagc ccgcggcagc tggcatccgcg gcagtccgcg gcagtccgcg gcagtccgcg tgggccttcg tgggcct tgggccttgggcct tgggcct tgggccttgggcct tgggcc tgggcct tgggcc tgggcc tgggcc tgggcc tgggcc tgggcc tggggcc tgggcc tggggcc tggggcc tggggcc tggggcc tggggcc tggggcc tggggcc tggggcc tggggcc tgggggggggg	* cgtcgccgcc ccTrcgCCCCCC * cgtcgccgcc ccTrcgCCCCCC * aaactcggcgcc ccTrcgCCCCCC * * aaactcgggc AAACTCCGCCC * tcgacaagga tcgacaggccagg cgaagtccag cgaagtccag cGAAGTCCAG	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcccecccaa gcaccccccaa gcacccccaa gcacccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcaccccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccccc	* actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggacgac t actcggacgacgac actcgacgacgacgacgacgacgacgacgacgacgacgacgac	* ttacctgtcci ttacctgtcci ttacctgtcc ttacctg	* agetegagee ggetegagee tggetegagee tggetegagee tggetegagee tggetegagee tggetegagege cgregegegege cgregegegege cgregegegegegegegegegegegegegegegegegege	* igtacgcagad igtacgcagad ogtacgcagad ogtacgcagad corracGCAGAG * igtacgcagad	* 2012 201
5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>qcqqaaa</u> 6001> <u>qcqqaaa</u> 6001> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u>	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * aggtgtactt: GGTGTACTTT * aaattcgac AAATTCGAC * ggggcggaga GGGCGGACA * acccgaaca acccgaaca	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt gcgtacgaga ggcgtacgaga gGCGTACGAGA * tcgagaagggc tcgagaagggc tcgagaagggc ctcacacggct catcaccggct	* ccgcggcagce ccGccGcGcAGC * ccgcggcagcagc ccGcCGCGCGCAGC * gcagtccgcg GCAGTCCGCG * tggcattcgcg GCAGTCCGCG * taggccttcg tgggccttcg tgggccttg tgggcctg tgggccttg tgggcctg tgggcctg tgggcg tgggcg tgggcg tggggg tgggg tggggc tgggg tggg tgggg tggg tgggg tgggg tgggg tggg t	* cgtcgccgcc cGTcGCCCCCCC * cgtcgccgcc cGTCGCCCCCCC * aaactcgggc AAACTCCGCCCCC * tcgacaagga tcgacaagga tcGACAAGGA * cgaagtccag cGAAGTCCAG	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcggcaccccaa gcaccccccaa gcacccccccccc	* actcggactac actggacggactac actggacggacgacac actggacggacgacac acggacggacgacac actggacgacacac actggacgacacacacacacacacacacacacacacacac	* ttacctgtcci ttacctgtcci TACCTGTCC' * ttacctgtcc ttacctgtcc ttacctgtcc gggatcgaca CGGATCGACA * tgggcgattc ggggcgattc GGGCGATTC * taggcctgatc CAGCCTGATC	* cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cgtcgagggggggggg	* gtacgcagac gtacgcagac GTACGCAGAC * gtacgcagac GTACGCAGAC * gtacgcagac CGTACGCAGAC * t gcacaccgaac GCACACCGAAC * gcccgcatcac gcccgcatcac gcccgcatcac gcccgcatcac gcccgcatcac gcccgcatcac gcccgcatcac gcccgcatcac	* cct>6000 cct>6000 cct>6000 cct>6000 cct>6000 cct>6000 cct>6100 cct
5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 24>TCCCCGA 6001> <u>qcqgaaa</u> 6001> <u>qcqgaaa</u> 6001> <u>qcqgaaa</u> 6101> <u>qcqgaaa</u> 6101> <u>qcqgaaa</u> 6101> <u>qcqgaaa</u> 6201> <u>qctqqat</u> 6201> <u>qctqqqt</u> 6201> <u>qctqqqt</u>	* ggtgtactt: GGTGTACTTV * ggtgtactt: GGTGTACTTV * aaattcgac .aaattcgac .aaattcgac .aaattcgac .aaattcgac .aaattcgac .aaattcgac .aaattcgac .aaatcgacaaaa GGGCGGAGAA *	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga ggcgtacgaga ggcgtacgaga tcgagaagggc tcgagaagggc tcgagaagggc tcgacacggct catcaccggct catcaccggct	* ccgcggcagc ccGCGGCAGCA * ccgcggcagc ccGCGGCAGCA * cgcagcagc gcagtccgcg gcagtccgcg gcAGTCCGCG * tgggccttcg tgggccttcg * accggggacct *	* categacaaca catagtcage categacaaca categacaaca aaactegage aaactegage aaactegage aaactegage aaactegage aaactegage categacaagga tegacaagga tegacaagga tegacaagga cagaagtecag cgaagtecag cGaagtecag	* gcgggcaccci gcGggcaccci gcGggcaccci gcGggcaccci gcGggcaccci gcGcGcCcAcCCi * gcacccccaa gcAcCCCCAA * tgccgcaaacci tgcCgcaaaci tgcCGCAAACi * tgccgcaaaci tgcCCCAACi *	* actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggacgacgac ggtcgggggggggg	* ttacctgtcci ttacctgtcci ttacctgtcc ttacctgtcc ttacctgtcc ttacctgtcc ttacctgtcc ttacctgtcc ttacctgtcc ttacctgtcc tacctgtcc tacctgttcc tacctgttcc tacctgttcc tacctgttcc tacctgttcc tacctgttcc tacctgttcc tacctgttcc tacctgttcc tac	* aggtcgagoc ggctcgagoc veGcrcGAGCC * tagctcgagoc reGcrcGAGCC * tagtcgagoc cgtcggcgg cgtcggggggg cgrcGGGCCGGGCCCG	* igtacgcagad igtacgcagad ogtacgcagad ogtacgcagad ig	* cct>6000 ccT>908 * cct>6000 ccT>223 * gac>6100 gac>6100 GAC>323 * agg=6200 agg=6200 agg=6200 agg=6200 agg=6200 ccT>523
5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 809>TCCCCGA 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>gcqqaaa</u> 224>GCGGAAA 6101> <u>gcqqcat</u> 6101> <u>gcqqcat</u> 6101> <u>gcqqcat</u> 6201> <u>gctqqt</u> 6201> <u>gctqqt</u> 6201> <u>gctqqt</u>	* ggtgtactt: GGTGTACTT * ggtgtactt: GGTGTACTT * aaattcgac AAATTCGAC * gggcggaga GGGCGGAGA * aacccgaaca acccgaaca ACCCGAACA	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt gcgtacgaga ggcgtacgaga gGcGTACGACA * tcgagaagggc tcgagaagggc tcgagaagggc tcgacgagaggc tcgacgagaggc tcgacgagggc tcgacgagggc tcgacgacggct cAtcAccGGCT *	* ccgcggcagce ccGccGGCAGCC * ccGccGGCAGCC * acgcggcagce gccAGCCGCGGCAGCC * tggccttcg tgggcettcg tgggcettcg * accggggacct * accggggacct * * * * * *	* cgtcgccgcc cGTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCCC * aaactcgggc AAACTCGGCC * tcgacaagga tcgacaagga tCGACAAGGA * cgaagtccag cGAAGTCCAG	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcGGGCCACCCI * gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa tgccgcaaac tgccgcaacc tgccgcaaac tgccgcaaac tgccgcaacc tgccgcaaac tgccgcaaac tgccgcaacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccgcacc tgccacc tgccacc tgccacc tgccacc tgccacc tgcc tgccgcacc tgccaccc tgccacc tgccacc tgccacc tgccacc tgccacc tgccacc tgccacc	* actcggactac actcggacgacgac actcggacgacgacgacgacgacgacgacgacgacgacgacga	* ttacctgtcci ttacctgtcci ttacctgtcci * ttacctgtcc * ttacctgtcc * ttacctgtcc * tgggatcgaca GGATCGACA * tgggcgattc GGGCGATTC * taggcctgatc aggcctgatc * taggcctgatc * taggcctgatc * * * * * * * * * * * * * * * * * * *	* cggctcgagcc cggctcgagcc tggctcgagcc tggctcgagcc cgccccGAGCC * tggctcgagcc cgccccGAGCC * cgtcgaggcgg cgtcggggggg cgrcGGGGGGCCCC * cgaggtggccc cGAGGTGGCCCC *	* gtacgcagac gtacgcagac GTACGCAGAC * gtacgcagac GGTACGCAGAC * gcacacccgaac CCACACCGAAC * gcccgcatcac gcccgcatcac gcccgcatcac * gcccgcattcgg CCGCATTCGGG *	* 201 201 201 201 201 201 201 201
5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>qcqqaaa</u> 224>CCCGGAA 6101> <u>qcqqcat</u> 6101> <u>qcqqct</u> 6101> <u>qcqqct</u> 6101> <u>qcqqct</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u> 6301> <u>qqtcqqc</u>	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * aaattcgac AAATTCGAC * gggcggaga GGGCGGAGA * acccgaaca acccgaaca ACCCGAACA * aggttacgag	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga gGcGTACGAGA * cagagaagggc tcgagaagggc tcgagaagggc catcaccggct catcaccggct catcaccggct catcaccggct catcaccggct catcaccggct catcaccggct catcaccggct	* ccgcggcagc ccGccgGcAGCC * ccGccGGCAGCC * ccGccGGCAGCC * gcagtccqcg GCAGTCCGCG * tgggcattcgcg GCAGTCCGCG * taggccttcg * accggggacct * accggggacct * accggggacct	* cgtcgccgcccccccccccccccccccccccccccccc	* gcgggcaccc: gcgggcaccc: gcgggcaccc: gcgggcaccc: gcgggcaccc: gcggcaccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa ccccacgcatact: ccccaacgtct: c	* actcggactac AcTcGGACTAC * actcggactac AcTcGGACTAC * actcggactac ACTCGGACTAC * actcggactac ACTCGGACTAC * agttcgcggtg GTTCGCGGTG GGTCGCGGCGCC * agtggcggcgcc CGGGCGGGCCCTAC * actggggggggcttc GGTGCGGCCTAC * cgacgccttcs *	* ttacctgtcci ttacctgtcci TTACCTGTCC' * ttacctgtcc ttacctgtcc TTACCTGTCC' * tgggatcgaca cggatcgaca cggatcgaca agggcgattc agggcgattc agggcgattc aggcctgatc cagcctgatc cagcctgatc * * * * * * * * * * * * * * * * * * *	* cggctcgagcc cggctcgagcc cgccccaccaccaccaccaccaccaccaccaccaccacca	* cgtacgcagac cGTACGCAGAG * cgtacgcagac cGTACGCAGAG * cgtacgcagac cGTACGCAGAG * cgtacgcagac cGTACGCAGAG * cgcacgcatcac cCACACCGAAG * cgcccgcatcac ccGCATCAG * cgccgcatcag ccgcattcagg ccgcattcagg ccgcattcagg ccgcattcagg ccgcattcagg ccgcattcagagagag	* cct>6000 ccT>908 * cct>6000 ccT>908 * cct>6000 cct>6000 ccT>223 * aggc>6100 gac>6100 gac>6100 gac>6200 aggs>6200 aggs>6200 acgc>6200 cct>6300 ccT>523 * aggc>6400
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> <u>qcqqaaa</u> 6001> <u>qcqqaaa</u> 224>GCGGAAA 6101> <u>qcqqca</u> 6101> <u>qcqqca</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u>	* ggtgtactt: GGTGTACTT * ggtgtactt: GGTGTACTT * aggtgtactt: GGTGTACTT * aaattcgac AAATTCGAC * gggcggaga GGCCCGAACA * acccggacaa ACCCGAACA * aggtacgaga	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga ggcgtacgaga gGcGTACGAGA * tcgagaagggc tcgagaagggc tcgagaagggc ctcaccggct catcaccggct catcaccggct catcaccggct caggtcgccgaa	* ccgcggcagca ccgcggcagca * ccgcggcagca ccgcgggcagc ccgccgcgcagca ccgccgcgcagca * gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg gcagtccgcg tgggcettcg * tgggcettcg * accgggacct * accgggacct * accgggacct * accgggacct	* cgtcgccgccgcc cgtcgccgccgcc cgtcgccgc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccgcc cgtcgccc cgtcgcc cgtcgcc cgtcgcc cgtcgccc cgtcgcccc cgtcgccccc cgtcgccccc cgtcgccccc cgtc	* gcgggcaccc; gcgggcaccc; gcgggcaccc; gcgggcaccc; gcgggcaccc; gcgcgccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcacccccaa gcaccccccaa gcaccccccaa gcacccccccccaa gcacccccccccc	* actcggactac actcgacgtgc actcgacgact actggacgacgac actcgacgacttca actcgacgacttca actcgacttca actcgact	* ttacctgtcci ttacctgtcci ttacctgtcc ttacctgccg	* cggctcgagcc cggctcgagcc cggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcgc cgcccccccc tggctcgaggggg cgrccGcGcGcCc tggagtggccc cGAGGTGGCCC	* gtacgcagad gtacgcagad ggtacgcagad ggtacgcagad ggtacgcagad cGTACGCAGAG * cacacccgaad cCACACCGAAG * gcccgcatcad ccccGCATCAG * ccgcattcagg	* 2012 2013 2014 2014 2014 2014 2014 2014 2014 2014
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001>qcqqaaa 224>CCCGAAA 6101>qcqqact 6101>qcqqct 324>CCCGCAAA 6101>qcqqct 6201>qctqqt 6201>qctqqt 6201>qctqqt 6301>qqtcqqc 534>GCCGCCCC	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * ggtgtgtactt: GGTGTACTTT * aaattcgac AAATTCGAC * gggcggaga GGGCGGAGA * acccgaaca acccgaaca ACCCGAACA * agttacgag	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga caggagaggc cagagagggc catcaccggct catcaccgc	* ccgcggcagc: ccGcGGGCAGCG * ccGcGGGCAGCG * ccGcGGGCAGCC * gcagtccqcg GCAGTCCGCG * tggcattcqcg GCAGTCCGCG * accgggaccttcg * accgggaccttcg * accgggaccttcg * accgggaccttcg * accgggaccttcg * accgggaccttcg * * accgggaccttcg * * accgggaccttcg * * accgggaccttcg * * accgggaccttcg * * accgggaccttcg * * accgggaccttcg * * accgggaccttcg * * accgggaccttcg * * * accgggaccttcg * * * * accgggaccttcg * * * * * * * * * * * * *	* cgtcgccgcc CGTCGCCCCCCC * cgtcgccgcc cGTCGCCCCCCC * aaactcgggc AAACTCCGGCC * cgaagtccagg cGAAGTCCAG * cgaagtccag gaactacgcg gaactacgcg gaactacgcg	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgcccccaa gcaccccgcac h	* actcggactac actcggacggacg gttcgcggtg gttcgcggtg ggtgcggggg actgggcttc ggtgcgggttt ggtgcggcttc actggacgcttc actggacgcttc actggacgccttc actggacgcctc actggacgcct actggacgacgcct actggacgacggacgacg actggacgacgacgacgacgacgacgacgacgacgacgacgac	* ttacctgtcci ttacctgtcci TACCTGTCC' * ttacctgtcc ttacctgtcc TTACCTGTCC' * tggatcgaca GGGATCGACA * agggcgattc GGGCGATTC * aggcctgatc CAGCCTGATCC * aggcctgatc aggcctgatc aggcctgatc aggcctgatc aggcctgatc * aggcctgat	* cggctcgagcc cggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggccgagcggc tggtcggggcgg tggtcggggcg tggtcggggcg tggtcggggcc tggtcggggcc tggtcggggcc tggtcggggcc tggtcggggcc tggtcgggcc tggtcggcc tggtcggcc tggtcggcc tggtcggcc tggtcggcc tggtcggcc tggtcgcc tggtcgcc tggtcgcc tggtcgcc tggtcgcc tggtcgccc tggtcgcc tggtcgcc tggtcgccc tggtcgccc tggtcgccc tggtcgccc tggtcgcccc tggtcgcccc tggtcgccccc tggtcgccccc tggtcgcccccc tggtcgcccccc tggtcgcccccc tggtcgcccccccc	* gtacgcagac GTACGCAGAC * gtacgcagac GTACGCAGAC * gtacgcagac GTACGCAGAC * t ggtacgcagac CGTACGCAGAC * t gccccgcatcac GCCCGCATCAG * t gcccgcattcgg CCGCATTCGG * t gcccggagagg CCGCATCGGAGAG	* 2015 2015 2015 2015 2015 2015 2015 2015
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901>tccccqa 124>TCCCCGA 6001>qcqqaaa 224>GCCGGAA 6101>qcqqcaa 224>GCCGGAA 6101>qcqqcaa 6201>qcqqca 6201>qcqcqa 6201>qcqqca 6201>qcqqca 6201>qcqqca 6201>qcqqca 6201>qcqqa 6201>qcqqca 6201>qcqqca 6201>qcqqa 6201>qcqqa 6201>qcqqca 6201>qcqqqa 6201>qcqqa 6201>qcqqa 6201>qcqqa 6201>qcqqa 6201>qcqq	* ggtgtactti GGTGTACTT * ggtgtactti GGTGTACTT * aaattcgac AAATTCGAC * gggcggaga GGCGGGAGAA * acccgaaca ACCCGAACA * * agttacgag agttacgag agttacgag	* ctccggctcgt cTCCGGCTCGT * ctccggctcgt ctccggctcgt gcgtacgaga ggcgtacgaga ggcgtacgaga ggcgtacgaga ggcgtacgaga ctcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc catcaccggct catcaccgct catcaccggct catcaccggct cat	* ccgcggcagc. ccGcGGCAGCC * ccgcggcagc. gcagtcagcagc gcagtcagcg gcAGTCCGCGG * tgggcattcg gGAGTCCGCG * tgggccttcg tgggccttcg accgggacct * accgggacct * accgggacct * acggtctcgac acggtctcgac acggtctcgac acgtctcgac	* cstcgccgccccccccccccccccccccccccccccccc	* gcgggcaccc; gcgggcaccc; gcgggcaccc; gcgggcaccc; gcgggcaccc; gccccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; gcacccccaa; tcccaacqtct; cccaacqtct; cccaacqtct; cactcggcqt; cactcgcqcd; cactcgcqcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcgcd; cactcd; cactcgcd;	* actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggacgac actcggacgacg actcggacgacgacgacgacgacgacgacgacgacgacgacga	* ttacctgtcci ttacctgtcci ttacctgtcc ttacctg	* agetegagee ggetegagee tggetegagee tggetegagee tggetegagee tggetegagee tggetegagege cgreggggggg cgreggggggg cgreggggggg cgaggtggee tggaatccccad A	* igtacgcagad igtacgagag igtacgagag igtacgagagag igtacgagag igtacgagagag igtacgagagag igtacgagagag igtacgagag igtacgagagag igtacgagag igtacgagag igtacgagagag igtacgagagag igtacgagagag igtacgagag igtacgagagag igtacgagag igtacgag igtacgagag igtacgagag igtacgagag igtacgagag	* 2012 2013 201
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> <u>qcqqaaa</u> 6001> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6101> <u>qcqqaaa</u> 6201> <u>qcqqaaa</u> 6201> <u>qctqqqt</u> 6201> <u>qctqqqt</u> 6201> <u>qctqqqt</u> 6201> <u>qctqqqt</u> 6201> <u>qctqqc</u> 6301> <u>qqtcqqc</u> 524>GCTCGCC	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * agattcgac * aaattcgac AAATTCGAC * * ggggcggaga GGGCGGACA * acccgaaca acccgaaca * ACCCGAACA * * agttacgag agtacgaga	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc tcgagaagggc ctcacaccggct catcaccggct catcaccggct catcaccggct catcaccggct catcaccggct catcaccggct catcaccggct catcaccggct catcaccgct	* ccgcggcagca ccgcggcagca ccgcggcagca ccgcggcagca ccgcggcagca ccgcggcagca ccgcggcagca ccgcggcagca gcagtccgcg GCAGTCCGCG * tgggccttcg tgggccttcg tgggccttcg accgggacct * accgggacct * accgggacct acgctccgac acgtctcgac	* cgtcgccgcc cGTCGCCCCCC * cgtcgccgcc cGTCGCCCCCC * cgctcgccgcc cGTCGCCCCCC * cgctcgccgcc cGTCGCCCCCC * cgaaactcgggc AAACTCCGGCC * cgaagtccag cGAAGTCCAG GGAACTACGCG	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcggcaccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccaa gcaccccccaa gcacccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccccccc	* actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggacgcgcgc actcgggcgggggggggg	* ttacctgtcci ttacctgtcci * ttacctgtcc * ttacctgtcc * ttacctgtcc * tgggatcgaca GGATCGACA * tgggcgattc GGGCGATTC * taggcctgatc CAGCCGATCC * taggcctgatc CAGCCGATCC * taggcctgatc CAGCCGATCC	* cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggctcgagcc cggcgggggggggg	* gtacgcagad ggtacgcagad GTACGCAGAG * ggtacgcagad GGTACGCAGAG * ggtacgcagad CGTACGCAGAG * ggcccgcatcad ggccgcatcad ggccgcatcad ggccgcatcad ccgcattcagg ccgcattcagg ccgcattcagagagag ccccGAGAGAG	* cct>6000 ccT>908 * cct>6000 cct>6000 cct>6000 ccT>223 * gac>6100 gac>6100 gac>6100 gac>6200 agg>6200 agg>6200 agg>6200 cct>6300 ccT>523 * gcg>6400 gcg>6400 gcg>6420 %
5901> <u>tccccqa</u> 5901>tccccqa 809>TCCCCGA 5901> <u>tccccqa</u> 5901>tccccqa 124>TCCCCGA 6001> <u>gcqgaaa</u> 224>GCCGAAA 6101> <u>gcqgcat</u> 6101>gcggcat 324>GCCGCCT 6201> <u>gctqqqt</u> 6201> <u>gcqqcqqt</u> 6201> <u>gctqqqt</u> 6201> <u>gcqqcqqt</u> 6201> <u>gcqqcqqt</u> 6201> <u>gcqqcqt</u> 6201> <u>gcqqc</u> 6201> <u>gcqcqc</u> 6201> <u>gcqc</u> 6201> <u>gcqcqc</u> 6201> <u>gcqc</u> 6201> <u>gcq</u> 6201> <u>gcq</u> 6201> <u>gcq</u> 6201> <u>gcq</u> 6201> <u>gcq</u> 6201> <u>gcq</u> 6201> <u></u>	* ggtgtactti ggtgtactti GGTGTACTT * iggtgtactt GGTGTACTT * aaattcgac aaattcgac aaattcgac AAATTCGAC * gggcggaga GGGCGGAGAA * acccgaaca ACCCGAACA * agttacgag AGTTACGAG *	* ctccggctcgt cTCCGGCTCGT * ctccggctcgt cTCCGGCTCGT * gcgtacgaga ggcgtacgaga gGCGTACGAGA * tcgagaagggc tcgagaagggc tcgagaagggc catcaccggct catcaccggt catcaccgg	* ccgcggcagc. ccGCGGCAGCC * ccgcggcagc. gcgcggcagc. gcgcggcagc. gcgcggcagc. tgggcatccgcg gcagtccgcg gcagtccgcg tgggcettcg tgggcettcg taccgggacct * accgggacct * accgggacct * accgggacct * tggtctcgac accgtctcgac acgtctcgac * * * * * * * * * * * * * * * * * * *	* cstcgccgcc ccrcgccccccc ccrcgcccccccc aaactcgggc aaactcgggc aaactcgggc tcgacaagga tcgacaagga tcgacaagga cgaagtccag gcaActacgcg gaactacgcg t	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcacccccaa gcaccccccaa gcacccccaa gcacccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccaa gcacccccaa gcacccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccccaa gcacccccccaa gcacccccccaa gcacccccccccc	* actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggactac actcggacgactac actcgacgacg gttcgcggtgc gggcgggcgc accGGCCGGCGCC * aggacggcgtcgc acGGCCGGCCCTC actgacgacttcc cGacgccttcc cGACGCCTTC/ *	* ttacctgtcc ttacctgtcc ttacctgtcc ttacctgtcc tacctgtcc tacctgttcc tacctgtttcc tacctgttcc tac	* aggtcgagoc gggctcgagoc gggctcgagoc tggctcgagoc tggctcgagoc tggctcgagoc tggctcgagoc tggctcgagoc tggctcgagoc tggtcggggggg cgrcgGccGGGCCGG * aggtgggcg cGrcGGGCCGG * aggtgggcg cGrcGGGCCCG * aggtggcg cGrcGGGCCCG * aggtggcg cGrcGGGCCCG * aggtggcg cGrcGGGCCCA * aggtggcg cGrcGGCCCA * aggtggcg * aggtgg * agg	* igtacgcagad igtacgagag igtacgagagg igtacgagag igtacgagagg igtacgagagg igtacgagagg igtacgagagg igtacgagagg igtacgagagg igtacgagagg igtacgagagg igtacgagagg igtacgagag igtacgag igtacgagag igtacgagag igtacgagag igtacgagag igtacgagag igtacgagag igtacgagag igtacgagag igtacgagag igtacgag igtacgagag igtacgagag igtacgag igtacgagag igtacgagag igtacgagag igtacgag igtacgagag igtacgagag igtacgagag igtacgagag igtacgag igtacgagag igtac	* 2012 2013 2014 2014 2014 2014 2014 2014 2014 2014
5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 124>TCCCCGA 6001> <u>qcqqaaa</u> 224>GCGGAAA 6101> <u>qcqqcat</u> 6101> <u>qcqqcat</u> 6201> <u>qctqqt</u> 6201> <u>qcqt</u> 6201> <u>qcqqt</u> 6201> <u>qcqqt</u> 6201> <u>qcqqt</u> 6201> <u>qcqt</u> 6201> <u>qcqqt</u> 6201> <u>qcqqt</u> 6201> <u>qcqt</u> 6201> <u>qcqqt</u> 6201> <u>qcq</u>	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * agattcgac aaattcgac AAATTCGAC * gggcggaga GGGCGGAGA * acccgaaca acccgaaca acccgaaca acccgaaca * CCCGAACA * * agttacgag CAGTTACGAG * * tcgggcgaaa	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt gcgtacgaga ggcgtacgaga gGcGTACGAGA * tcgagaagggc tcgagaagggc tcgagaagggc tcgacgaga ggcgtacggaga ctcaccggct catcaccggct catcaccgccga caggtcgccga	* ccgcggcagce ccGcGGGCAGCC * ccGcGGGCAGCC * acagtccgcg gccActccgcg gccActccgcg tgggcettcg tgggcettcg tgggcettcg taccggggcact * accggggacct * accggggacct * acgtctcgac * acgtctgac * acgtcgac * acgtctgac * acgtct	* cgtcgccgcc cGTCGCCGCCC * cgtcgccgcc cGTCGCCGCCCC * aaactcgggc AAACTCGGGC * tcgacaagga tcgacaagga tcGAAAGTCCAGG cgaagtccag gaactacgcgg GAACTACGCGC *	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcGGGCCACCCI * gcacccccaa gcAcccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcaccccccaa gcacccccaa	* actcggactac agtgcgggggg agtgcgggcttc agtgcggcttc acgacgccttcc acgacgccttcc acgacgactac actggacgactac actgg	* ttacctgtcc ttacctgtcc * ttacctgccg * ttactcgccg * ttacctgccg * ttactcgccg * ttactcgccg * ttacctgccg * ttactcgccg * ttacctgccg * ttactcgccg * ttactgccg * ttactgccg * ttactgccg * ttactgccg * ttactcgccg * ttactcgcc	* cggctcgagcc cggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctggccc tggctgggcgg cgtcggggggg cgtcggggggg cgtcggggggg cgtcggggggg cgtcggggggg cgtcggggcc tggaggtggcc tggagttggcc tggagttggcc tggagttggcc tggagttgccc tggaatccccac	* statespagae st	* 2015 2015 2015 2015 2015 2015 2015 2015
5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 809>TCCCCGA 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 5901> <u>tccccqa</u> 24>TCCCCGA 6001> <u>qcqqaaa</u> 6001> <u>qcqqaaa</u> 224>GCGGAAA 6101> <u>qcqqct</u> 6101> <u>qcqqct</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u> 6201> <u>qctqqt</u> 6301> <u>qqtcqqc</u> 524>GCCGCGC	* ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * ggtgtactt: GGTGTACTTT * aaattcgac aaattcgac AAATTCGAC * ggggcggaga GGGCGGACA * acccgaaca accccgaaca acccgaaca a	* ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ctccggctcgt ggcgtacgaga caggtagcgcga catcaccggct cat	* ccgcggcagce ccGccGGCAGCC * ccGccGGCAGCC * ccGccGGCAGCC * gcagtcagcg gcAgtcagcg GCAGTCCGCG * tgggcattcagcg GCAGTCCGGCA * tagggcettcg tagggcettcg * accgggaacct * accgggaacct * accgggacct * accggcacca	* cgtcgccgcc cGTCGCCGCCCC * cgtcgccgcc cGTCGCCGCCCCC * cgtcgccgcc cGTCGCCGCCCCC * aaactcgggc AAACTCGGCC * tcgacaagga tcgacaagga tcgacaagga cGAAGTCCAG * cgaagtccag ggaactacgcg ggaactacgcg gGAACTACCCG	* gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcgggcaccci gcggcaccccaai gcacccccaacgcct cccaacgcct caccccgcctcc cacccctcccc	* actcggactac actcggacggg actac actcggacgg actac actcggacgacge actggacggacge actggacgacgac actggacgacgacgac actggacgacgac actggacgacgacg	* ttacctgtcci TACCTGTCC' * ttacctgtcc TTACCTGTCC' * ttacctgtcc TTACCTGTCC' * tgggtcgaca GGGATCGACA * agggcgattc GGGCGATTC * aggcctgatc AGGCCGATTC * aggcctgatc AGGCCGATCC * aggcctgatc aggcctgatc * aggc	* cggctcgagcc cggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggctcgagcc tggtcggggcgg cgtcgggggggg cgtcgggggggg	* gtacgcagac GTACGCAGAG * gtacgcagac GTACGCAGAG * gtacgcagac GTACGCAGAG * tacccgaa CCACACCGAA * gcccgcatcas GCCGCATCAG * gcccgcatcagagag CCCGGAGAGC * GCCGCATCGGAGAGC * GCCGCACACCGAGAGAGC * GCCGCACACCGAGGAGAGC * GCCGCACCGAGGAGAGC * GCCGCACCGAGGAGAGC * GCCGCACCGAGGAGAGC * GCCGCACCGAGGAGGAGC * GCCGCGCACCGAGGAGG * GCCGCGCACCGAGGAGG * GCCGCGCACCGAGGAGG * GCCGCGCACCGAGGAGG * GCCGCGCACCGAGGAGG * GCCGCGCACCGCAC	* 2015 2015 2015 2015 2015 2015 2015 2015

Fig. A3-2. Alignment of sequencing result showing successful construction of recombinant plasmid pET11a-ssuD. T7 promoter starts at 5230 bp. ssuD insert starts at 5321 bp and ends at 6476 bp. Two primers were used in sequencing: pET11a-f (5'-ATA CCC ACG CCG AAA CAA GC -3') and ssuD-f (5'-TAA AGT CGA TCA CGA CGA CC-3').

E201>30300	*	*	*	*	*	*	*	*	*	*
	TCGAGATCTC	GATCCCGCG	AAATTAATAC	GACTCACTAT	AGGGGAATTO	TGAGCGGATA	CAATTCCCC	TCTAGAAATA	ATTTTGTTT	ACTTT>5300
5201>ACACC	MCCACAMOMO	CATCCCCCC	A A M M A A M A A	CACTOACTAT	ACCCCA AMMO	TCACCCATA	CAAMMCCCCC	TOTACAAAT	Ammmomm	ACTTT 5300
107- AGAGG	TCGAGATCIC	GATCCCGCG	AAATTAATAC	GACICACIAI	AGGGGAATTC	TGAGCGGATA		TCIAGAAAIA	ATTTGTT	ACTIT>5500
10/>AGAGG	ATCGAGATCTC	GATCCCGCG	AAATTAATAC	GACTCACTAT	AGGGGAATTC	TGAGCGGATA	ACAATTCCCCC	TCTAGAAATA	ATTTTGTTTA	ACTTT>206
	*	*	*	*	*	*	*	*	*	*
E201>88088		mamcatata			+ ant anna a		a a b a a a b a a	a a a a b a a a a b		
5301-ANGAN	SCAGATATACA	TAIGatyte	acayaccaac	guuduuguu		Cerdededed	lecteggiea	accyccayct	.cyccyayacc	gegae 5400
5301>AAGAA	GAGATATACA	TATGatgtc	acagaccaac	gttctcgttc	tcgtcggcag	cctgcgcgccc	gcctcggtca	accgtcagct	cgccgagaco	gcgac>5400
207>AAGAA	GAGATATACA	TATGATGTC	ACAGACCAAC	GTTCTCGTTC	TCGTCGGCAC	CCTGCGCGCCC	SCCTCGGTCA	ACCGTCAGCI	CGCCGAGACO	CCCGAC>306
	*	*	*	*	*	*	*	*	*	*
E 4015										
5401>cgccg	cggcgcccgcg	ggtgccgag	gtgaccgtgt	tcgacggcct	cggagacgto	cccttctaca	icgaggacat	cgacgtcgcc	ggttccatco	ccgcc>5500
5401>cgccg	radedeecded	ggtgccgag	gtgaccgtgt	tcgacggcct	cggagacgto	cccttctaca	acgaggacat	cgacgtcgcc	ggttccatco	ccgcc>5500
307>CGCCG	regegeceecg	GGTGCCGAG	GTGACCGTGT	TCGACGGCCI	CGGAGACGTC	CCCTTCTACA	ACGAGGACAT	CGACGTCGCC	GGTTCCATCO	CCGCC>406
CC015-1										
5501>gtcga	gcgctgcgtg	aggccgccg	gccgcgccga	cgcgctcctg	cttctcacgo	ccgagtacaa	ggcaccatc	cccgccgtcc	tcaagaacgo	catcg>5600
5501>gtcga	gcgctgcgtg	aggccgccg	gccgcgccga	cgcgctcctg	cttctcacgo	ccgagtacaa	cggcaccatc	cccgccgtcc	ctcaagaacgo	catcg>5600
407>GTCGA	GCGCTGCGTG	AGGCCGCCG	GCCGCGCCGA	CGCGCTCCTC	CTTCTCACGO	CCGAGTACAA	CGGCACCATC	CCCGCCGTCC	TCAAGAACGC	CATCG>506
							-			
5601>actgg	tctcccggcc	gtacggcac	cggagccgtg	aaggacaago	ccgtcgccgt	catcagegee	cgeceageg	gcaacggtgc	acagtggggcg	cacga>5700
5601>actgg	atctcccggcc	at pagage pa	aaaaaaata	22442422244	aaakaaaak	cot coacaca				$a_{0} = a_{0} = 5 = 700$
		gcacggcac	cygayccycy	aayyacaayo	cegeegeege	cattagegee	regeccageg	gcaacggtgc	cacagtgggcg	cacya>5700
507>ACTGG	TCTCCCGGCC	GTACGGCAC	CGGAGCCGTG	AAGGACAAGO	CCGTCGCCG	CATCAGCGCC	CGCCCAGCG	GCAACGGTGC	CACAGTGGGGC	CACGA>606
507>ACTGG	ATCTCCCGGCC	GTACGGCAC	CGGAGCCGTG	AAGGACAAGC	CCGTCGCCG	CATCAGCGCC	CGCCCAGCG	GCAACGGTGC	CACAGTGGGGCG	CACGA>606
507>ACTGG	ATCTCCCGGCC	GTACGGCAC	CGGAGCCGTG	AAGGACAAGC	CCGTCGCCG	CATCAGCGCC	CGCCCAGCG	GCAACGGTGC	CACAGTGGGGCG	CACGA>606
507>ACTĞĞ	ATCTCCCGGCC	GTACGGCAC	CGGAGCCGTG	AAGGACAAGC	CCGTCGCCG	CATCAGCGCC	CGCCCAGCG	GCAACGGTGC	acagtgggcg ACAGTGGGCG	CACGA>606
507>ACTGG	ATCTCCCGGCC	GTACGGCAC	CGGAGCCGTG	AAGGACAAGC	CCGTCGCCG	*	CGCCCAGCG	GCAACGGTGC	*	*
507>ACTGG	ATCTCCCGGCC *	gtccgcatc	cGGAGCCGTG * gccggcggca	aggacaago AAGGACAAGO * aggttctcga	*	*	*	Cgacaagtto	*	*
507>ACTGG 5701>ggaca 5701>ggaca	ATCTCCCGGCC * cccgcaaggcc	gtccgcatc gtccgcatc	cGGAGCCGTG * gccggcggca gccggcggca	aggacaago AAGGACAAGO * aggttotoga	* CCGTCGCCGT Aggacgtcacc	* cATCAGCGCCC * ctcgcatcgc	* rcgcccAgcg * rcggcaccat	cgacaagttc	ACAGTGGGGG ACAGTGGGGG * :ggtgaccgg :ggtgaccggg	* * * * * * * * * * * * * *
507>ACTGG 5701>ggaca 5701>ggaca 607>GGACA	ATCTCCCGGCC * cccgcaaggcc cccgcaaggcc cccgcaaggcc	gtacggeac gtacggeac gtccgcatc gtccgcatc	* gccggcggcg gccggcggca gccggcggca	AAGGACAAGC * aggttctcga aggttctcga AGGTTCTCGA	* * * * * * * * * * * * * *	* cATCAGCGCCC * ctcgccatcgc ctcgccatcgc cTCGCCATCGC	* rcgcccAgcg rcggcaccat rcggcaccat rcggcaccat	cgacaagtto	ACAGTGGGCG * CACAGTGGGCG * CGGTGACCGGC CGGTGACCGGC	* * * * * * * * * * * * * *
507>ACTGG 5701>ggaca 5701>ggaca 607>GGACA	ATCTCCCGGCC * cccgcaaggcc cccgcaaggcc cccgcAAGGCC	tacggeac GTACGGCAC <u>stccgcatc</u> gtccgcatc GTCCGCATC	¢ gccggcggca gccggcggca gccggcggca	* aggttctcga aggttctcga AGGTTCTCGA	* * * * * * * * * * * * * *	* cATCAGCGCC * ctcgccatcgc ctcgccatcgc cTCGCCATCGC	* rcgcccAgCG rcggcaccat rcggcaccat rcggcaccat	cgacaagtto cgacaagtto cgacaagtto cGACAAGTTO	ACAGTGGGGG CACAGTGGGGGG * cggtgaccggg cGGTGACCGGG	* * * * * * * * * * * * * *
507>ACTGG 5701>ggaca 5701>ggaca 607>GGACA	ATCTCCCGGCC * cccgcaaggcc cccgcaaggcc cccgcAAggcC	gtacggeate gtacgcatc gtccgcatc gtccgcatc GTCCGCATC	cGGAGCCGTG * gccggcggca gccggcggca GCCGGCGGCA	AAGGACAAGC * aggttctcga AGGTTCTCGA	* tggacgtcacc tggacgtcacc tggacgtcacc tggacgtcacc	* * * * * * * * * * * * * * * * * * *	* acggcaccat acggcaccat acggcaccat acggcaccat	GCAACGGTGC GCAACGGTGC * cgacaagttc CGACAAGTTC	ACAGTGGGGG CACAGTGGGGGG * cggtgaccggg cGGTGACCGGG	* :acccq>5800 :acccq>5800 :ACCCG>706
507>ACTGG 5701>ggaca 5701>ggaca 607>GGACA	ATCTCCCGGCC * CCCGCAAGGCC CCCGCAAGGCC *	gtacggaac GTACGGCAC gtccgcatc gtccgcatc GTCCGCATC	cGGAGCCGTG * <u>qccqqcqqca</u> gccggcggca GCCGGCGGCA *	AAGGACAAGC * aggttctcga AGGTTCTCGA	teggeegtegeegt * teggacgtcacc teggacgtcacc tegacgtcacc	* * :ctcgccatcgg :ctcgccatcgg :ctcgccatcgg :ctcgccatcgg	* gcggcaccat gcggcaccat GCGGCACCAT	cgacaagttc cgacaagttc cGACAAGTTC	acagtggggg chChGTGGGCC * ggtgaccggg GGTGACCGGC *	* * * * * * * * * * * * * *
507>ACTGG	* * * * * * * * * * * * * *	gtacggeac gtacgGCAC <u>gtccgcatc</u> gtccgcatc GTCCGCATC	cGGAGCCGTG * <u>accqqcqqca</u> gccqgcggca GCCGGCGGCA	AAGGACAAGC * aggttctcga AGGTTCTCGA	* aggacgtcacc Ggacgtcacc GGACGTCACC	* * * * * * * * * * * * * *	* <u>acggcaccat</u> <u>acggcaccat</u> <u>acggcaccat</u> <u>acggcaccat</u> <u>acgcACCAT</u>	cgacaagttc cgacaagttc cgacaagttc	* * * * * * * * * * * * * *	* * * * * * * * * * * * * *
507>ACTGG 5701>ggaca 5701>ggaca 607>GGACA 5801>cgcga	* * * * * * * * * * * * * *	gtacggcac gtacggcatc gtccgcatc gtccgcatc gtccGCATC *	cGGAGCCGTG * <u>accggcggca</u> gccggcggca GCCGGCGGCA * <u>aggtcgccga</u>	AAGGACAAGC * aggttctcga aggttctcga AGGTTCTCGA * cgtcgtcgcc	teggeggeggeggeggeggeggeggeggeggeggeggegg	* catcagecatege ctcgccatege ctcgccatege ctcgccatege ctcgccatege *	rcgcccAgCg rcgcccAgCg rcggcaccat rcggcaccat rcggcAcCAT rcggctcgtcg rcgctcgtcg	cgacaagttc cgacaagttc cgacaagttc cGACAAGTTC * gacgcctgaQ	ACAGTGGGGGG * cagtgaccggc cggtgaccggc cGGTGACCGGC * CGTGACCGGC	* * * * * * * * * * * * * *
507>ACTGG 5701>ggaca 5701>ggaca 607>GGACA 5801>cgcga 5801>cgcga	TCTCCCGGCC * CCCGCAAGGCC CCCGCAAGGCC * TAACGCCGAGG yaacgccgagg	gtacggeac gtacggeate gtccgcatc gtccgcatc gtccgcatc tcgcgcagc	cGGAGCCGTG * gccggcggca gccggcggca gcCGGCGCCA * aggtcgccga aggtcgccga	AAGGACAAGC * aggttetega aggttetega AGGTTETEGA * cgtegtegecgec	Aggacgtcacc s aggacgtcacc ag	* ctcgccatcgc ctcgccatcgc ctcgccatcgc ctcgccatcgc ctcgccatcga acgccacgaa acgccacgaa	* rcGCCCAGCG * rcggcaccat rcggcaccat scGGCACCAT * rcagctcgtc rcgcagctcgtc	cgacaagttc cgacaagttc cgacaagttc cGACAAGTTC <u>*</u> gacgcctgaC gacgcctgaC	ACAGTGGGGCG * cagtgaccggc cggtgaccggc cGTGACCGGC * ccraccarcac ccraccarcac	* * * * * * * * * * * * * *

Fig. A3-3. Alignment of sequencing result showing successful construction of recombinant plasmid pET11a-ssuE. T7 promoter starts at 5230 bp. ssuE insert starts at 5321 bp and ends at 5885 bp. One primer was used in sequencing: pET11a-f (5'-ATA CCC ACG CCG AAA CAA GC -3').