

**KARADENİZ TECHNICAL UNIVERSITY  
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**

**DEPARTMENT OF BIOLOGY**

**MODIFICATION OF HYDROGEN METABOLISM FOR IMPROVED ETHANOL  
PRODUCTION IN THERMOPHILIC BACTERIA**

**Ph.D. THESIS**

**Ayşenur EMİNOĞLU M. Sc.**

**JANUARY 2018**

**TRABZON**



**KARADENİZ TECHNICAL UNIVERSITY**  
**THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES**



**This thesis is accepted to give the degree of**

**By**  
**The Graduate School of Natural and Applied Sciences at**  
**Karadeniz Technical University**

**The Date of Submission** : / /

**The Date of Examination** : / /

**Thesis Supervisor** :

**Trabzon**

KARADENİZ TECHNICAL UNIVERSITY  
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

BIOLOGY  
AYŞENUR EMİNOĞLU M. Sc.

MODIFICATION OF HYDROGEN METABOLISM FOR IMPROVED ETHANOL  
PRODUCTION IN THERMOPHILIC BACTERIA

Has been accepted as a thesis of

DOCTOR OF PHILOSOPHY

after the Examination by the Jury Assigned by the Administrative Board of the  
Graduate School of Natural and Applied Sciences with the Decision Number 1735 dated

09 / 01 / 2018

Approved By

Chairman : Prof. Dr. Hikmet GEÇKİL  
Member : Prof. Dr. Ali Osman BELDÜZ  
Member : Prof. Dr. Yusuf TUTAR  
Member : Prof. Dr. Sabriye ÇANAKÇI  
Member : Asst. Prof. Kadriye İNAN BEKTAŞ

  
.....  
.....  
.....  
.....  
.....

Prof. Dr. Sadettin KORKMAZ  
Director of Graduate School

## FOREWORD

The study presented in this thesis was performed at the Prof. Lee R. LYND Biofuel Research Lab, Department of the Chemical and Biochemical Engineering, Thayer School of Engineering Dartmouth.

This study was supported by TUBITAK-BIDEB International Doctoral Research Fellowship Program (2214/A) with a scholarship and by The BioEnergy Science Center (which is an United States Department of Energy (DOE) Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science) with Dartmouth College under United States Energy contract no. DE-AC05-00OR22725.

I would like to express my sincere gratitude to my advisor Prof. Dr. Ali Osman BELDÜZ for his guidance and for his support to me and to my ideas during my P.hD studies and to Prof. Lee R. LYND for accepting me as a visiting researcher to his magnificent biofuel research group and giving me the opportunity to work with the most important scientists in their field, it was a great pleasure to be a member of the Lynd Lab. I also would like to express my great thanks to Dr. Daniel G. OLSON for his invaluable scientific guidance and help during this study. Also, my special thanks go to Sean Jean-Loup MURPHY, Marybeth MALONEY, Dr. Anthony LANAHAN, Dr. Evert HOLWERDA and to all Lynd Research Group for their help and support during the research. I would also like to express my special thanks to Prof. Dr. Sabriye ÇANAKÇI, Ass. Prof. Kadriye İNAN BEKTAŞ and to all KTU Molecular Biology Research Lab members, and also many great thanks to my colleagues at Recep Tayyip Erdogan University, Department of Biology. I am also very grateful to my friends, Esmâ CEYLAN and Ass. Prof. Ayşe Hümeyra T. KAFA for their moral support.

This doctoral thesis is a result of working hard in good faith and believing that Allah's help will always be with me no matter how hard the path may be. I am grateful to my brothers and to my sister for their support throughout my life. I dedicate this thesis to my departed father Dr. (MD) Numan Fazıl EMİNOĞLU and to my mother Birnaz EMİNOĞLU who has been sacrificed so much for her kids throughout her life.

Ayşenur EMİNOĞLU

Trabzon 2018

## **STATEMENT OF ETHICS**

I declare that all the parts of this Ph. D. thesis entitled “Modification of Hydrogen Metabolism for Improved Ethanol Production in Thermophilic Bacteria” has been accomplished under the supervision of my advisor Prof. Dr. Ali Osman BELDÜZ. I also declare that I have collected the data myself, I have completed the experiments/analysis in the related laboratories and I observed the highest ethical and scientific standards during the development of this thesis. I avoided plagiarism and fully acknowledged the work of others to which I have referred in this thesis. I admit all the legal consequences if the falsehood of this declaration is proven. 29/01/2018

## TABLE OF CONTENTS

	<u>Page No</u>
FOREWORD.....	III
TABLE OF CONTENTS.....	V
ÖZET.....	IX
LIST OF FIGURES .....	XI
LIST OF TABLES .....	XVI
LIST OF ABBREVIATIONS .....	XVII
1. GENERAL INFORMATION .....	1
1.1. Introduction .....	1
1.2. Biomass as an Energy Resource.....	5
1.3. Lignocellulosic Biomass .....	7
1.3.1. Structure of Lignocellulose .....	8
1.4. Converting Lignocellulosic Biomass to Biofuels.....	12
1.5. Ethanol Production From Lignocellulose.....	13
1.6. Consolidated Bioprocessing (CBP).....	15
1.7. Metabolic Engineering of Thermophilic Bacteria For Ethanol Production .....	19
1.7.1. Metabolic Engineering of <i>T. saccharolyticum</i> for Ethanol Production.....	20
1.7.2. Metabolic Engineering of <i>C. thermocellum</i> for Ethanol Production.....	23
1.8. Fe-Only [FeFe] Hydrogenases .....	26
1.9. Biofuel in Turkey .....	27
2. MATERIALS AND METHODS .....	28
2.1. Chemicals and Kits.....	28
2.2. Buffers and Stock Solutions .....	29
2.4. Bioinformatic Analysis and Determination of the Microorganisms to be Studied .....	32
2.5. Primers.....	32
2.6. Molecular Studies.....	39
2.6.1. Construction of <i>hfsA, B, C, D</i> Knockout <i>Thermoanaerobacterium</i> and <i>Thermoanaerobacter</i> Strains .....	39
2.6.1.1. Preparation of Gene Deletion Cassettes .....	39
2.6.1.2. Transformation of the Deletion Fragments .....	43

2.6.2.	Deletion of the <i>hfsB</i> gene in <i>Clostridium thermocellum</i> DSM 1313 .....	44
2.6.2.1.	Construction of the Gene Deletion Plasmid .....	44
2.6.2.2.	Transformation of <i>hfsB</i> -pDGO145 into LL345 via Electroporation.....	46
2.6.2.3.	Colony Selection and Marker Removal .....	47
2.6.3.	Deletion of <i>hfsAB</i> and <i>hfsCD</i> in <i>Thermoanaerobacterium saccharolyticum</i> Strain.....	48
2.6.4.	Deletion of Lactate Dehydrogenase ( <i>ldh</i> ) Gene in <i>T. saccharolyticum</i> $\Delta$ <i>hfsA</i> (LL1267) and $\Delta$ <i>hfsB</i> (LL1268) Mutant Strains.....	49
2.6.5.	Whole Genome Sequence Analysis .....	50
2.6.5.1.	Bacterial Genomic DNA Isolation With CTAB.....	50
2.6.5.2.	Whole Genome Sequence Analysis .....	52
2.6.6.	Differential Expression Analysis of Mutant <i>T. saccharolyticum</i> Strains by RNA Sequencing .....	52
2.6.7.	Reverse Transcription Quantitative PCR (RT-qPCR) Expression Analysis to Determining the Expression Levels .....	52
2.6.8.	Quantitative Proteomics Analysis of Mutant Strains by MS/MS.....	53
2.6.9.	Heterologous Expression of <i>hfsABCD</i> * Gene in <i>Clostridium thermocellum</i> Under Three Different Promoter.....	54
2.6.9.1.	Plasmid and Strain Constructions.....	54
2.6.9.2.	Transformation of the Expression Plasmids into <i>Clostridium thermocellum</i> $\Delta$ <i>hydG<math>\Delta</math><i>ech</i> .....</i>	57
2.6.9.3.	Overexpressing the <i>hfsABCD</i> operon in $\Delta$ <i>hydG<math>\Delta</math><i>ech</i> <i>C. thermocellum</i> (LL1147).....</i>	57
2.7.	Sample Collection and Analytical Studies .....	58
2.7.1.	Bottle Fermentation and Analysis of the Fermentation Products.....	58
2.7.1.1.	pH Meseurements.....	58
2.7.1.2.	Pressure.....	58
2.7.1.3.	Meseurement of H <sub>2</sub> and CO <sub>2</sub> Percentage .....	58
2.7.1.4.	HPLC and TOCN (Total Organic Carbon) Analysis .....	59
2.7.2.	Fermentation Conditions for <i>C. thermocellum</i> strains .....	59
2.8.	Enzyme Assays.....	60
2.8.1.	Preparation of Cell Free Extract.....	60
2.8.2.	Alcohol Dehydrogenase (ADH) - Aldehyde Dehydrogenase (ALDH) Assays ....	60
2.8.3.	Hydrogenase Assays.....	61
3.	RESULTS.....	62

3.1.	Deletion of the <i>hfsA</i> , <i>hfsB</i> , <i>hfsC</i> and <i>hfsD</i> Genes in <i>T. saccharolyticum</i> with <i>KanR</i> replacement.....	63
3.2.	Deletion of <i>hfsB</i> Gene in <i>T. thermosaccharolyticum</i> , <i>T. xylanolyticum</i> , and <i>T. mathranii</i> .....	66
3.3.	Removal of <i>hfsB</i> in LL345 ( <i>C. thermocellum</i> $\Delta$ <i>hpt</i> ) .....	68
3.4.	Deletion of <i>hfsAB</i> and <i>hfsCD</i> Clusters in <i>T. saccharolyticum</i> .....	74
3.5.	Inactivation of Lactate Dehydrogenase ( <i>ldh</i> ) in LL1267 and LL1268 .....	77
3.6.	Enzyme Assays.....	81
3.7.	RNAseq and Proteomics Data Analysis .....	83
3.8.	RT-qPCR Expression Analysis .....	86
3.9.	Whole Genome Resequencing Analysis.....	87
3.10.	Fermentation Product Distribution of <i>T. saccharolyticum</i> Strains and Analysis ...	88
3.11.	Analysis of <i>T. thermosaccharolyticum</i> , <i>T. xylanolyticum</i> and <i>T. mathranii</i> Fermentation Product Distribution .....	92
3.12.	Heterologous Expression of <i>hfsABCD</i> * Gene in <i>Clostridium thermocellum</i> Under Three Different Promoter .....	93
3.13.	Fermentation Profile of Overexpression of <i>hfsABCD</i> Operon in <i>C. thermocellum</i> $\Delta$ <i>hydG</i> $\Delta$ <i>ech</i> (LL1147) .....	96
3.14.	Deletion of <i>hfsB</i> in $\Delta$ <i>hpt</i> <i>C. thermocellum</i> DSM 1314.....	97
4.	DISCUSSION .....	99
5.	CONCLUSION .....	103
6.	RECOMMENDATIONS .....	104
7.	REFERENCES .....	105
	CURRICULUM VITAE	



Ph.D. Thesis

SUMMARY

MODIFICATION OF HYDROGEN METABOLISM FOR IMPROVED ETHANOL  
PRODUCTION IN THERMOPHILIC BACTERIA

Ayşenur EMİNOĞLU  
Karadeniz Technical University  
The Graduate School of Natural and Applied Sciences  
Biology Graduate Program  
Supervisor: Prof.Dr. Ali Osman BELDÜZ  
2018, 119 Pages

In this study; the changes in the production of ethanol and hydrogen resulting from the cellobiose fermentation were evaluated comparatively by performing deletion of the *hfsA*, *hfsB*, *hfsC* and *hfsD* genes (one-by-one) of the *hfsABCD* operon in wild-type *Thermoanaerobacterium saccharolyticum* and it has been found that the deletion of *hfsA* and/or *hfsB* gene resulted in about 1.8 fold increase in mutant *T. saccharolyticum* strain when compared to the wild-type. Enzyme activity assays have shown that alcohol dehydrogenase (AdhE) activity in mutant strains, where ethanol production is high, is about 2.0 fold higher than the wild type. Transcriptomic and proteomic studies have shown that this increase in ethanol production is related, in partly, to an increase in the expression of the *adhE* gene, suggesting the existence of a regulatory mechanism between *hfsB* and *adhE*. It has also been shown that the deletion of the *hfsB* gene causes an increase in ethanol production in other thermophilic anaerobes containing *hfs*-like hydrogenase. Among them, *Thermoanaerobacterium xylanolyticum* and *Thermoanaerobacterium thermosaccharolyticum* mutants have the highest ethanol yields reported in the literature to date. This suggests that *hfsB* gene deletion can be used as a general genetic engineering method with the aim of increasing ethanol production in other microorganisms containing *hfs*-like hydrogenase.

**Key Words:** *Thermoanaerobacterium*, *Clostridium*, *Thermoanaerobacter*, Metabolic engineering, Hydrogenase, Ethanol.

Doktora Tezi  
ÖZET

TERMOFİLİK BAKTERİLERDE ETANOL ÜRETİMİNİN ARTIRILMASI AMACI İLE  
HİDROJEN METABOLİZMASININ MODİFİKASYONU

Ayşenur EMİNOĞLU  
Karadeniz Teknik Üniversitesi  
Fen Bilimleri Enstitüsü  
Biyoloji Anabilim Dalı  
Danışman: Prof. Dr. Ali Osman BELDÜZ  
2018, 119 Sayfa

Lignoselülozik biyokütle, doğada en fazla bulunan organik materyaldir. Bol miktarda bulunan lignoselülozik biyokütlenin etanol gibi sıvı yakıtlara dönüştürülmesi enerji güvenliğinin artırılması, ticaret açığının azaltılması, sera gazları emisyonunun düşürülmesi ve fiyat kararlılığının sağlanmasında önemli bir seçenek sunmaktadır. Konsolide Biyoproses (KBP) lignoselülozun enzim eklenmeksizin tek bir adımda istenilen ürünlere dönüştürülmesidir. Bu yöntem özellikle lignoselüloz karbohidratlarını doğal parçalama kapasitesine sahip mikroorganizmaların genetik olarak modifiye edilmeleri ile zorlu endüstriyel süreçlere dayanıklı, yüksek verim ve derişimde yakıt üretebilen mikroorganizmalara dönüştürülmelerine olanak sağlaması açısından son yıllarda giderek artan bir ilgiyle karşılanan bir araştırma sahasına dönüşmüştür.

Bu amaçla çok çalışılan organizmaların başında *Thermoanaerobacterium saccharolyticum* ve *Clostridium thermocellum* gelmektedir. *T. saccharolyticum*; ksilanı ve biyokütle türevli şekerlerin büyük bir çoğunluğunu yüksek sıcaklık derecelerinde fermente edebilen, hemiselüloolitik anaerob termofillerden bir tanesidir. *C. thermocellum* ise, *T. saccharolyticum*'un aksine fermentasyon için kristalize selülozu parçalayabilen ancak beş C'lu şekerleri kullanamayan selüloolitik bir anaerob termofildir. Tanımlanan tüm termofilik sakkarolitik anareoblar, fermentasyon sonucunda etanolün yanı sıra asetat, laktat, amino asit, CO<sub>2</sub>, H<sub>2</sub> ve başlıca fermentasyon ürünlerini üretmektedirler. *T. saccharolyticum* ve *C. thermocellum* gibi termofillerde asetik asit pirüvat ferrodoksin oksidoredüktaz (PFOR), fosfat asetil transferaz ve asetat kinaz vasıtası ile pirüvattan meydana gelirken, laktik asit L-

laktat dehidrojenaz vasıtası ile oluşmaktadır. Pirüvatın, PFOR vasıtası ile asetil-CoA'ya dönüşümü sonucu indirgenmiş ferrodoksin üretilir. Ferrodoksinde gelen elektronlar hidrojen üretiminde kullanıldıkları gibi etanol üretiminde de kullanılabilir ve teorik olarak hidrojen üretimindeki bir düşüşün etanol üretiminde eş değerde bir artışa, aynı şekilde hidrojen üretiminde bir artışın etanol üretiminde eş değerde bir düşüşe neden olması beklenir.

Yapılan bu çalışmada; yaban tipi olan *T. saccharolyticum* 'da *hfsABCD* operonunu oluşturan *hfsA*, *hfsB*, *hfsC* ve *hfsD* genlerinin teker teker delesyonu yapılarak sellobioz fermentasyonu sonucu etanol ve hidrojen üretiminde meydana gelen değişiklikler karşılaştırmalı olarak değerlendirilmiş ve *hfsA* ve/veya *hfsB* gen delesyonları sonucu *T. saccharolyticum*'da etanol üretiminin yaban tipe oranla yaklaşık 1.8 kat artış gösterdiği tespit edilmiştir. Enzim aktivite deneyleri etanol üretiminin yüksek olduğu mutant suşlarda alkol dehidrojenaz (AdhE) aktivitesinin yaban tipe göre yaklaşık 2.0 kat artış gösterdiğini ortaya koymuştur. Transkriptomik ve proteomik çalışmalar etanol üretiminde ki bu artışın, kısmen de olsa *adhE* geninin ekspresyonunda bir artışla ilişkili olduğunu göstermiş ve bu da *hfsB* ve *adhE* arasında düzenleyici bir mekanizmanın varlığını düşündürmüştür. Ayrıca *hfsB* gen delesyonunun *hfs*-benzeri hidrojenaz içeren, çalışılan diğer termofilik anaeroblarda da etanol üretiminde artışa neden olduğu ortaya konmuştur. Bunlar içerisinde *Thermoanaerobacterium xylanolyticum* ve *Thermoanaerobacterium thermosaccharolyticum* mutant suşlarından literatürde bugüne kadar bildirilen en yüksek etanol verimi elde edilmiştir. Bu da *hfsB* gen delesyonunun *hfs*-benzeri hidrojenaz içeren diğer mikroorganizmalarda da etanol üretiminin artırılması amacı ile genel bir genetik mühendisliği yöntemi olarak kullanılabileceğini ortaya koymuştur.

*hfsB* geninde nokta mutasyonları taşıyan *T. saccharolyticum* suşu (LL1186) *hfsABCD* operonunun (*hfs\**),  $\Delta hydG\Delta ech$  *C. thermocellum* suşunda (LL1147) üç farklı promotor altında heterolog ekspresyonu sonucu elde edilen mutant suşlarda ise kontrol suşa nazaran etanol üretiminde herhangi bir artış elde edilemediği görülmüştür.

**Anahtar Kelimeler:** *Thermoanaerobacterium*, *Clostridium*, *Thermoanaerobacter*, Metabolik mühendislik, Hidrojenaz, Etanol.

## LIST OF FIGURES

	<u>Page No</u>
Figure 1. Comparison of cost and ethanol production quantities of CBP and simultaneous saccharification and fermentation (SSCF) with enzymatic hydrolysis [5].....	2
Figure 2. The chemical structure of cellulose [47].....	9
Figure 3. Structure of hemicellulose [50].....	10
Figure 4. Representation of lignin structure [52].....	11
Figure 5. Schematic illustration of lignocellulose [42].....	12
Figure 6. Bioethanol production process steps, SHF: separate hydrolysis and fermentation; SSF: simultaneous saccharification and fermentation; SSCF: simultaneous saccharification and co-fermentation; CBP: consolidated bioprocessing; C6: hexose; C5: pentose [3].....	17
Figure 7. A schematic representation of the products formed after pyruvate in glucose degradation in obligative anaerobes, ALDH: acetaldehyde dehydrogenase; ADH: alcohol dehydrogenase; AK: acetate kinase; FNOR: ferredoxin oxidoreductase; H <sub>2</sub> -ase: hydrogenase; LDH: lactate dehydrogenase; PFOR: pyruvate ferredoxin oxidoreductase; PTA: phosphotransacetylase [3].....	24
Figure 8. Schematic illustration of ethanol production pathway in thermophiles [15].....	25
Figure 9. Map of the backbone vector pDGO145 was used for <i>hfsB</i> disruption in <i>C. thermocellum</i> .....	46
Figure 10. Map of the <i>hfsABCD</i> expression vector in <i>C. thermocellum</i> .....	55
Figure 11. <i>hfsABCD</i> operon (4.86 kb) chromosomal organization and domain structure scheme in <i>T. saccharolyticum</i> .....	62
Figure 12. <i>hfs A, B, C</i> and <i>D</i> deletion fragments generated by PCR; L: DNA Ladder (NEB-N0469S); 1: <i>hfsA</i> 3' flank; 2: <i>hfsA KanR</i> ; 3: <i>hfsA</i> 5' flank; 4: <i>hfsB</i> 3' flank; 5: <i>hfsB KanR</i> ; 6: <i>hfsB</i> 3' flank; 7: <i>hfsC</i> 5' flank; 8: <i>hfsC KanR</i> ; 9: <i>hfsC</i> 3' flank; 10: <i>hfsD</i> 5' flank; 11: <i>hfsD KanR</i> ; 12: <i>hfsD</i> 3' flank.....	63
Figure 13. <i>hfsA, B, C</i> and <i>D</i> deletion cassettes which were verified and generated by PCR from assembled fragments; L: DNA ladder (2-log quick load purple-NEB); A: <i>hfsA</i> replacement cassette (AE2-AE6), B: <i>hfsB</i> replacement cassette (AE7-AE11), C: <i>hfsC</i> knockout cassette (AE12-AE16) and D: <i>hfsD</i> replacement cassette (AE17-AE21).....	64
Figure 14. Confirmation of <i>hfsA</i> deletion by colony PCR in selected colonies; L: 1 kb DNA ladder (N3232L); 1, 2, 3, 4, 5, 6, 7 and 8 are selected colonies, wt: wild-type, 1.5-2.0 kb bands show amplification performed with primers targeting the internal region of the <i>KanR</i> gene.....	64
Figure 15. Confirmation of <i>hfsB</i> deletion by colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); 1, 2, 3, 4, 5, 6, 7 and 8 are selected colonies, wt: wild type, 1.5-2.0 kb bands show	

amplification performed with primers targeting the internal region of the <i>KanR</i> gene.....	65
Figure 16. Confirmation of <i>hfsC</i> deletion by colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); 1, 2, 3, 4, 5, 6, 7 and 8 are selected colonies, wt: wild-type, 1.5-2.0 kb bands show amplification performed with primers targeting the internal region of the <i>KanR</i> gene.....	65
Figure 17. Confirmation of <i>hfsD</i> deletion by colony PCR in selected colonies; L:1 kb DNA ladder (N3232L); 1, 2, 3, 4, 5 are selected colonies, wt: wild-type, 1.5-2.0 kb bands show amplification performed with primers targeting the internal region of the <i>KanR</i> gene.....	66
Figure 18. PCR amplified <i>hfsB</i> deletion fragments for transformation; L:1 kb DNA ladder (N3232L); 1: for <i>T. xylanolyticum</i> ; 2: for <i>T. thermosaccharolyticum</i> ; 3: for <i>T. mathranii</i> .....	67
Figure 19. <i>hfsB</i> deletion confirmation in <i>T. xylanolyticum</i> and <i>T. thermosaccharolyticum</i> ; L:1 kb DNA ladder (N3232L); 1: internal control in LL1347; 2: internal control in LL1301; 3: external control in LL1347; 4: external control in LL1301; 5: Internal control in LL1346; 2: Internal control in LL1244; 3: external control in LL1346; 4: external control in LL1244.....	67
Figure 20. <i>hfsB</i> gene deletion internal confirmation in <i>T. mathranii</i> ; L: 1 kb DNA ladder (N3232L); 1, 2, 3, 4, 5, 6, 7 were selected colonies; Wt: LL1258.....	68
Figure 21. Fragments of <i>hfsB</i> gene deletion in <i>C. thermocellum</i> , 5'Flnk-Int 3'Flnk: flanking regions for homologous recombination, V1-V2: pDGO145 vector backbones; L: 1 kb DNA ladder (N3232L).....	69
Figure 22. Internal region clonin confirmation in <i>E. coli</i> clones with primes AE119-AE120, the band size ~1.0 kb selected as positive clone; L: 1 kb DNA ladder (N3232L).....	69
Figure 23. Confirmation of cloning the 5' and 3' flanking regions in <i>E. coli</i> clones with primers AE115-AE118; L: 1 kb DNA ladder (N3232L); the band size ~2.5 (5'flnk+3'flnk) kb selected as positive clone.....	70
Figure 24. Confirmation of integrity of selection markers in selected <i>E. coli</i> clones with primers AE198-AE199 ( <i>cat-hpt</i> : the band size 616 bp) and AE200-AE201 ( <i>cbp-tdk</i> : the band size 1029 bp); (+) vector control; L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively.....	71
Figure 25. Confirmation of integrity of selection markers <i>cbp-tdk</i> in selected <i>C. thermocellum</i> clones on Tm-FUDR with the expected size of 1029 bp (primers AE200-AE201); L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively.....	71
Figure 26. Confirmation of integrity of selection markers <i>cat-hpt</i> in selected <i>C. thermocellum</i> clones on Tm-FUDR with the expected size of 616 bp (primers AE198-AE199); L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively.....	72
Figure 27. Confirmation of last step of <i>hfsB</i> gene deletion after selection with 8-AZH with the expected band size of 2366 bp (primers AE222-AE223); L: 1 kb	

DNA ladder (N3232L).....	72
Figure 28. Confirmation of <i>hfsB</i> gene deletion in selected <i>C. thermocellum</i> clones after last step; L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively.....	73
Figure 29. Confirmation of removal of selection markers <i>cat-hpt</i> and <i>cbp-tdk</i> in selected <i>C. thermocellum</i> clones after last step; L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively.....	73
Figure 30. <i>hfsAB</i> and <i>hfsCD</i> knockout cassette; L: 1 kb DNA ladder (N3232L); <i>hfsAB</i> and <i>hfsCD</i> knockout cassettes amplified by PCR including kanamycin resistance gene assembled with up and down regions (Up- <i>KanR</i> -Down).....	74
Figure 31. $\Delta hfsAB::kan$ colony PCR bands with the external primers AE24 and AE27, L: quick load@2-log DNA ladder (N0469S).....	75
Figure 32. $\Delta hfsAB::kan$ colony PCR bands with the primers AE24 (external) and AE23 (internal), L: quick load@2-log DNA ladder (N0469S).....	75
Figure 33. $\Delta hfsCD::kan$ colony PCR bands with the external primers AE28 and AE31; L: 1 kb DNA ladder (N3232L).....	76
Figure 34. $\Delta hfsCD::kan$ colony PCR bands with the primers AE28 (external) and AE23 (internal); L: quick load@2-log DNA ladder (N0469S).....	77
Figure 35. <i>ldh</i> knockout fragments generated by PCR, L: 1 kb DNA ladder (N3232L); <i>EryR</i> ; erythromycin resistance cassette; Up; <i>ldh</i> 5' flanking fragment Dwn; <i>ldh</i> 3' flanking fragment.....	78
Figure 36. The <i>ldh</i> knockout cassette; L: 1 kb DNA ladder (N3232L); line 1; <i>ldh</i> knockout cassette including erythromycin resistance gene assembled with <i>ldh</i> Up and Down regions (Up- <i>EryR</i> -Dw).....	78
Figure 37. External confirmation of <i>ldh</i> knockout by two external primers via colony PCR in selected colonies, L: molecular weight marker (1 kb DNA ladder, N3232L); lines 1, 2, 3, 4, 5, 6, 7, 8 and 9 show the ~4.0 kb fragment of <i>ldh</i> deletion locus in selected colonies, $\Delta hfsA$ : <i>ldh</i> gene locus amplification fragment in parent strain by external primers with ~3.9 bp expected size.....	79
Figure 38. Internal confirmation of <i>ldh</i> knockout by colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); lines 1, 2, 6, 7, 8 and 9 show the ~1.9 kb internal fragment in selected colonies; $\Delta hfsA$ : negative control for erythromycin cassette.....	79
Figure 39. External confirmation of <i>ldh</i> knockout by colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); lines 1, 2, 4, 6, 7, 8, 9 and 10 show the ~4.0 kb external fragment of <i>ldh</i> deletion locus in selected colonies; $\Delta hfsB$ : <i>ldh</i> gene locus amplification result in parent strain by external primers with ~3.9 bp expected size.....	80
Figure 40. Internal confirmation of <i>ldh</i> knockout by colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); lines 1,2,4,6,7,8 and 9 show the ~1.9 kb internal fragment in selected colonies; $\Delta hfsB$ : negative control for erythromycin cassette.....	80

Figure 41. ALDH and ADH activities in <i>T. saccharolyticum</i> strains with the NADH cofactor, (SD=1; n ≥ 4).....	81
Figure 42. ALDH and ADH activities in <i>T. saccharolyticum</i> strains with the NADPH cofactor, (SD=1; n ≥ 4).....	82
Figure 43. BV reduction in <i>T. saccharolyticum</i> strains to test hydrogenase activity, (SD=1; n ≥ 3).....	83
Figure 44. Gene expression analysis of high (LL1267, LL1268 and LL1049) and wild-type level (wt <i>T. saccharolyticum</i> , LL1269, LL1270 and LL1187) ethanol-producing strains, the plot is showing difference in gene expression, genes with high expression in the high ethanol producing strains are plotted on the right side, the vertical axis is representing the significance of the change.....	84
Figure 45. <i>hfsA, B, C</i> and <i>D</i> genes expression analysis of the mutant strains, the <i>hfsA</i> deletion downregulates <i>hfsB</i> , which may explain its phenotype, the other <i>hfs</i> deletions have almost no effect on downstream gene expression, upregulation of <i>adhE</i> is observed by RT-qPCR, which agrees with RNAseq data.....	85
Figure 46. Plot graph of protein abundance vs wt (LC/MS-MS data), genes with 5-fold expression difference compared to wt are indicated with gene name or locus number, <i>adhE</i> is consistently over-expressed in high ethanol producing strains.....	86
Figure 47. Results of RT-qPCR analysis of <i>adhA</i> and <i>adhE</i> genes in <i>T. saccharolyticum</i> strains normalized to <i>recA</i> , (SD=1; n=2).....	87
Figure 48. Fermentation product distribution of various <i>T. saccharolyticum</i> strains from 5 g/L cellobiose, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n ≥ 2).....	88
Figure 49. Ethanol yield of several <i>T. saccharolyticum</i> strains grew on 5 g/ L cellobiose, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n ≥ 2).....	89
Figure 50. Fermentation product yield of LL1349 and LL1350, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n=3).....	90
Figure 51. Fermentation product yield of LL1349 and LL1350, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n ≥ 2).....	90
Figure 52. Fermentation product yield of LL1349 and LL1350, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n=3).....	91
Figure 53. Fragments of <i>hfsABCD</i> and promoter regions, L: DNA ladder; a: <i>hfsABCD</i> amplified by <i>eno</i> overlapping primers; b: <i>hfsABCD</i> amplified by 2638 overlapping primers; c: <i>hfsABCD</i> with native promoter region, <i>eno</i> and 2638 are promoter fragments.....	92
Figure 54. Linearized pDGO143 with <i>SmaI</i> .....	93
Figure 55. pDGO143-P <sub><i>eno</i></sub> - <i>hfsABCD</i> * colony PCR with an external and internal primer; L: DNA ladder; V: vector control.....	94

Figure 56. pDGO143- <i>P<sub>hfs</sub>-hfsABCD*</i> colony PCR with an external and internal primer; L: DNA ladder; V: vector control.....	94
Figure 57. pDGO143- <i>P<sub>2638</sub>-hfsABCD*</i> colony PCR with an external and internal primer, L: DNA ladder; V: vector control.....	95
Figure 58. Colony PCR results of the LL1147 clones with an external primer pair, which are containing the expression vectors; L: DNA ladder; lines 1, 2, 3: pDGO143- <i>P<sub>hfs</sub>-hfsABCD*</i> ; lines 4, 5, 6: pDGO143- <i>P<sub>2638</sub>-hfsABCD*</i> and lines 7, 8, 9: pDGO143- <i>P<sub>eno</sub>-hfsABCD*</i> ; V: vector control.....	96
Figure 59. Fermentation products of several mutant <i>C. thermocellum</i> strains, the parent strain <i>ΔhydGΔech</i> (LL1147) and expression strains were grown on MTC-5 with 5 g/L cellobiose, product yield in units of moles product, produced per mole of cellobiose consumed.....	97
Figure 60. Fermentation products of several mutant <i>C. thermocellum</i> strains that grown on MTC-5 with 5 g/L cellobiose, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n=3).....	98
Figure 61. Hypothetical mode of action of <i>hfs</i> operon on ethanol production.....	100



## LIST OF TABLES

	<u>Page No</u>
Table 1. Plasmids and strains used in this study.....	29
Table 2. List of primers used in this study .....	32
Table 3. PCR reaction conditions for knockout fragment amplification.....	40
Table 4. PCR conditions for <i>C. thermocellum hfsB</i> deletion fragment amplification.....	44
Table 5. PCR reaction conditions for <i>hfsAB</i> and <i>CD</i> assembly fragment amplification.....	48
Table 6. PCR reaction conditions for <i>ldh</i> assembly fragments. ....	50
Table 7. Description of promoter sequences used for <i>hfs</i> expression in <i>C. thermocellum</i> . 55	
Table 8. Fragment cycling conditions for <i>C. thermocellum hfsABCD</i> expression plasmid.56	

## LIST OF ABBREVIATIONS

8AZH	8-azahypoxanthine
<i>adhA</i>	alcohol dehydrogenase A
<i>adhE</i>	alcohol dehydrogenase E
CBP	Consolidated bioprocesses
CTAB	Cetyltrimethylammonium bromide
DMSO	Dimethyl sulfoxide
<i>eps</i>	exopolysaccharide
<i>erm</i>	Erythromycin
FBP	fructose 1,6 bisphosphate
FUDR	5-fluoro-2'-deoxyuridine
<i>hfs</i>	hydrogen-Fe-S hydrogenase
<i>kan</i>	Kanamycin
<i>ldh</i>	lactate dehydrogenase
PAS	<i>Per-Arnt-Sim</i> domain
PDC	Pyruvate decarboxylase
PDH	Pyruvate dehydrogenase
PFL	Pyruvate formate lyase
PFOR	Pyruvate ferredoxin dehydrogenase
<i>pta-ack</i>	phosphotransacetylase-acetylkinase
RT	Room temperature
T <sub>m</sub>	Thiempheicol

## 1. GENERAL INFORMATION

### 1.1. Introduction

Increasing depletion and consumption of oil and fossil fuel resources has played a prominent role in the development and use of alternative liquid fuels such as ethanol produced from lignocellulosic biomass [1]. The European Commission aims to meet at least 20% of energy production from second generation renewable sources by 2020 [2].

The vast majority of bioethanol produced to date has been produced from simple substrates such as starch (corn) and sucrose (sugar beet and sugar cane), known as the first generation. At the same time, however, the use of these sources, which are also used in food production, for biofuels has led to an increase in the price of food plants and anxiety about the consumption of food resources. However, the production of alternative fuels, called second generation that produced by lignocellulosic biomass, is a response to these concerns.

Today, conversion of lignocellulosic biomass into biofuels is accomplished through the use of industrial microorganisms such as yeast, either by the addition of cellulolytic enzymes or by heterologous production [3].

Alternatively, lignocellulosic carbohydrates can be transformed into biofuels by a microorganism that can able to produce ethanol with high yield and concentration, at the challenging industrial processes with genetic modifications. [4]. Today, for first-generation fuel production mostly mesophilic ethanologenic microorganisms such as *Saccharomyces cerevisiae* and *Zymomonas mobilis* which produces ethanol at high yield are used. Although these organisms have properties such as ethanol productivity at near theoretical values, high ethanol tolerance, they are ineffective in producing ethanol from lignocellulosic biomass due to limited substrate selectivity in wild strains [3].

Although cellulosic plant biomass has many desirable properties as an energy source, it is difficult to convert this valuable substance into liquid fuels. In this conversion process, biological processing technology is a promising option but still presents significant challenges. These difficulties stem from the fact that plants have developed to be resistant to attacks, especially to the microorganisms and their enzymes. In principle, this resistance barrier can be overcome in two ways: (1) with the innovation of thermochemical technologies, a field based on process engineering, involving processes in which

intermediates other than sugars (synthesis gas, pyrolysis etc) are separated, or (2) fermentative processes (acid hydrolysis, phosphoric acid swelling, pretreatment with ionic liquids) by way of innovation of biotechnology. Later, consolidated bioprocesses (CBP) is emerging which lignocellulosic biomass can be transformed into the desired products in a single step without enzyme addition. CBP can also be defined as the ultimate regulation of hydrolysis and fermentation of cellulosic biomass at low cost. CBP provides significant potential for obtaining products with significantly lower cost and much higher efficiency than all methods based on processing technologies based on cellulase production [5]. Over the past few years, intensive efforts have been made to reduce the cost of cellulase production [6]. According to the last report, the cost of producing cellulosic ethanol is about 2.5 to 5 cents per liter (10-20 cents per gallon ( $\text{¢} / \text{gal}$ )). However, the estimated cost with CBP is 1.25 cents per liter (4.23 cents per gallon) [5]. These features of CBP demonstrate the possibility of a more efficient production with much lower capital.

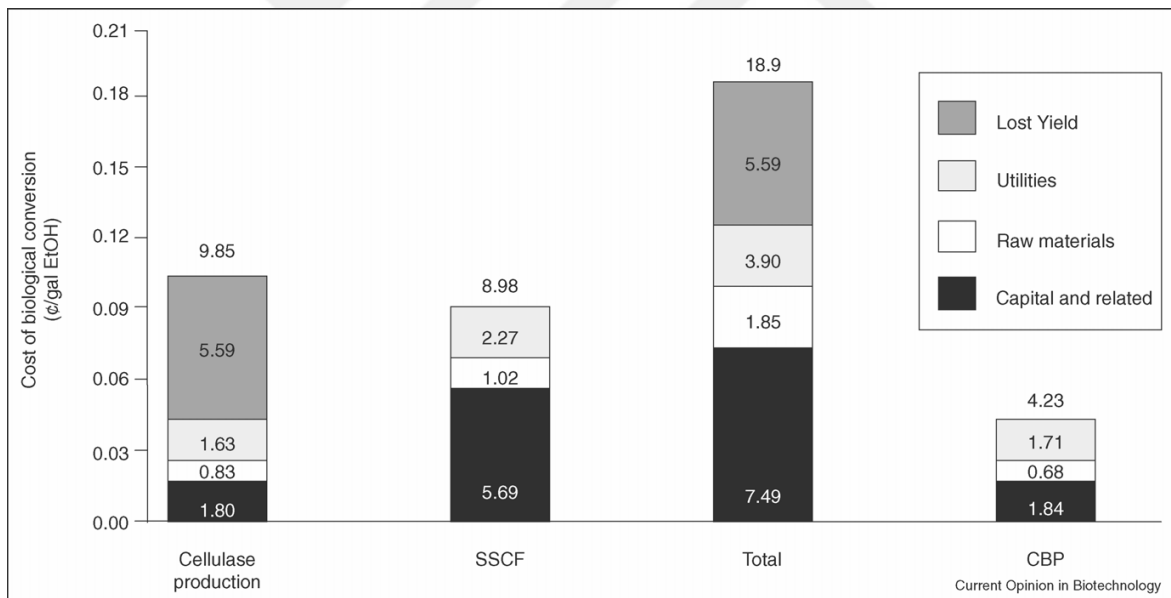


Figure 1. Comparison of cost and EtOH production quantities of CBP and simultaneous saccharification and fermentation (SSCF) with enzymatic hydrolysis [5].

This method has become a field of increasing interest in recent years, because of allowing the microorganisms that capable of hydrolysis of lignocellulosic carbohydrates, to be transformed into genetically modified microorganisms which are resistant to the harsh industrial processes that produce fuel at high yield and concentration. A microorganism to be used for CBP must break down a biomass substrate in industrial conditions and produce

the desired products in high yield and concentration. Since microorganisms with these properties are not found naturally, genetic engineering methods are needed in obtaining these organisms [7].

The reasons for this include a complex enzyme system known as a "cellulosome" that breaks down a variety of substrates such as cellulose, starch, xylenes, and even lignin, which naturally lie in the lignocellulosic biomass. Complex enzyme systems exhibit synergistic activity on lignocellulosic biomass much more efficiently and effectively than an enzyme alone can [8]. Among these anaerobic bacteria identified as etanologically, there are mainly *Clostridium*, *Caldanaerobacterium*, *Thermoanaerobacter* and *Thermoanaerobacterium* genus members [7]. As a result of studies with *Clostridium thermocellum* belonging to *Clostridium* genus and *Thermoanaerobacterium saccharolyticum* belonging to genus *Thermoanaerobacterium*, many techniques for applying gene transfer systems and genetic engineering in these bacteria have been developed and successfully applied [9, 10, 11].

The conversion of lignocellulose into liquid fuels in developed and developing countries has become an increasing demand because of its benefits of using available, renewable sources of cellulosic plant biomass for producing biofuels [12, 13, 14].

Among the studies carried out in this direction, the development of genetically modified microorganisms for CBP is at the center of the researches [5]. However, because of the recalcitrant nature of the lignocellulose to the foreign agents like microbial enzymes, producing biofuels in the manner of microbes at high yield and titer under industrial conditions needs some modification in the targeted microorganism metabolism via metabolic engineering. To obtain industrially useful microbes with desired features, the microorganism needs to engineer of more than one metabolic pathway in combination, to prevent undesired byproducts or growth deficiencies associated with the cause of environmental adaptation difficulties of the mutant strains [7, 8, 17, 18].

*Clostridium thermocellum*, a cellulase-forming anaerobic bacterium, is one of the most developed hosts for CBP applications. This is due to the high cellulose degradability of *C. thermocellum*. Recent studies on the metabolic properties of *C. thermocellum* [18] and advances in techniques developed for genetic modification [19, 20, 21] have shown that *C. thermocellum*, as a CBP organism, has become an attractive to obtain biofuels or biochemical products [22].

The members of the *Thermoanaerobacterium* has a great importance in ethanol production by metabolic engineering for CBP because of their effortless genetic engineering

with well-developed tools [23, 24, 25]. From this genus, the bacteria *Thermoanaerobacterium saccharolyticum*, *Thermoanaerobacterium xylanolyticum* and *Thermoanaerobacterium thermosaccharolyticum* from this genus were known to have native ability to utilize a variety of sugars in hemicellulose to produce ethanol [26, 27, 28]. Besides ethanol these microbes also producing organic acids (acetate, lactate, fumarate, etc.) and H<sub>2</sub> by fermentation [8, 17, 28]. In particular, there were some approaches has been applied in order to eliminate other products like lactic acids and acetic acids than ethanol in fermentation mixture and increase ethanol production in *T. saccharolyticum* [16, 29, 30]. Previously Saw et al. [31] generated an “ethanologenic” phenotype by deleting the genes for lactate and acetate production however they observed that the phenotype did not appear immediately, but only after a short period of adaptation. When those strains have been resequenced, it was observed that they had mutations in both the *adhE* (alcohol dehydrogenase) gene and the *hfs* (hydrogenase-Fe-S) hydrogenase. Attempts to reconstitute the ethanologenic phenotype by re-introducing the *adhE* or *hfs* point mutations were only moderately successful. In all of these methods, in order to obtain an industrial high ethanol producing phenotype that does not form by-products or forms only very low amounts of undesirable products, at least two or more gene should knocked-out or engineered. Thus an ethanologenic strain can only be obtained at the end of very complicated and time-consuming genetic works [32].

*T. saccharolyticum* was recently evaluated to produce ethanol in 92% theoretical yield with the elimination of the *Tsac\_0795*, *pta-ack*, *ldh* and *eps* [32]. However, there was no study has been found so far that resulted in with an ethanologenic phenotype strains of *T. thermosaccharolyticum* and *T. xylanolyticum* in the literature.

The study presented in this thesis aims to understand the role of the *hfs* operon and reveal the function and regulatory mechanisms of *hfsA*, *hfsB*, *hfsC* and *hfsD* genes involved in this operon. In the light of findings obtained from *T. saccharolyticum*, it is also investigated if the determined regulation system is present in other organisms having this operon and whether this method could be used as a metabolic engineering strategy to generate ethanologenic strains.

## 1.2. Biomass as an Energy Resource

Biomass has been the most important source of almost all basic needs throughout human history [33].

Biomass is a term used for all plant organic matter including algae, trees and all kind of crops. Biomass is produced by green plants that can transform inorganic matter into plant material through photosynthesis by taking advantage of sunlight, comprising all land and water-based vegetation including organic wastes. Organic materials in which sunlight energy is stored in chemical bond, is the source of biomass. The chemical energy stored in organic substances is released when the bonds between adjacent carbon, hydrogen, and oxygen molecules are broken by decomposition, combustion or digestion and used as a major source of the energy since the earliest times of mankind existence.

The value of a specific type of biomass depends on the chemical and physical properties of large molecules that constituted about it. Mankind has benefited from this stored energy in the chemical bonds by burning the biomass as fuel and by eating the plants for their nutrients of sugar and starch. Fossilized biomass is used as coal and oil. However, the conversion of biomass into fossil fuels takes millions of years and thus could not be renewed in a period that mankind could use. With the burning of fossil fuels, the old biomass is converted to CO<sub>2</sub>, which causes the greenhouse effect, and non-renewable resources are consumed. However, burning of the new biomass does not cause new CO<sub>2</sub> emissions to the atmosphere; because re-planting of the harvested biomass ensures that CO<sub>2</sub> is absorbed and reintroduced into the new growth cycle [34].

As a result of rising energy demand, the high cost of fossil fuels, diminishing of the fossil fuel reserves and impact of the fossil fuel usage on the greenhouse effect, biomass has become increasingly important globally as a clean energy source alternative to fossil fuels [34].

In addition to this, technological developments in the industrial field for the conversion of biomass has lowered the conversion cost and has led to higher yield. Developments in the agricultural sector in countries such as western Europe and the United States has also increased interest in the use of bio-energy. Biomass is also very important because it is an indigenous energy source, and it is available in many countries and its application may contribute to energy security by diversifying the fuel supply [34].

Since biomass resources being potentially the largest and most sustainable source of energy in the world, biomass has taken its place in almost all major global energy supply scenarios [33].

The European Commission has set a long-term goal by 2050 to develop a competitive, resource efficient low carbon economy and the green economy concept is included in the overall framework at different levels of EU policy.

This policy aims to improve the economy by reducing polluting emissions, increasing resource efficiency, preventing the loss of biodiversity and increasing the value of economic services [37]. Biomass energy originates from a variety of sources, which are classified as forests, agriculture, and wastes. Crops for biofuel, energy-producing grasses, short-rotation trees, woody biomass and residues, herbaceous by-products and municipal solid wastes are among potential sources. Globally, the largest share of biomass energy in 2012 was obtained from forests with a total supply of 49 EJ out of the total supply of 56.2 EJ. Today's present energy supply is about 560 EJ [37].

It is envisaged that the energy available from biomass will meet approximately 50% of the primary global energy need by 2050. Some of these biomass energy will be used for alcohol and diesel fuel production, the majority for combustible gas, and the remaining will be used for fuel power stations [38].

Biomass can be transformed into three main products: electricity heat energy, fuel for transportation and chemical feedstock [34]. Among others, fuel production from biomass is advantageous for two reasons. The first of these is the reduction of the net CO<sub>2</sub> emissions (greenhouse gas) given to the atmosphere. Unlike fossil fuels, biofuels are natural carbon fuels since the resulting CO<sub>2</sub> from combustion of the fuel is used for biomass growth immediately afterward. Furthermore, the establishment of a new biomass-based economic system will ensure that the economies in the respective countries are less sensitive to fluctuations in oil prices and will open up new business areas in areas such as agriculture, forestry, oil, and chemical industry. However, the use of edible biomass (sugar, starch and vegetable oils) for large-scale production brings with it a number of ethical questions. These questions have led many researchers to develop new technologies for the processing of non-edible biomass (lignocellulosic biomass). Thus, it is aimed to provide sustainable production of fuel (second-generation fuels) and chemicals without affecting existing food resources. In addition, because of the lower cost and rapid growth of lignocellulosic sources of biomass



when compared with food crops and its large availability makes them an attractive raw material that could be substituted by petroleum [39].

### **1.3. Lignocellulosic Biomass**

Increasing depletion of petroleum and fossil fuel resources and their increased consumption have led to a prominent role in the development and use of alternative liquid fuels such as ethanol, the raw material of which is renewable lignocellulosic biomass. Converting abundant lignocellulosic biomass to liquid fuels such as ethanol offers an important option to increase energy security, reduce trade barriers, reduce greenhouse gas emissions, and ensure price stability [1].

The chemical properties of their components make lignocellulose a biotechnologically valuable substrate [39, 40]. Forestry and agricultural activities, the paper-pulp and timber industry and many agricultural industries result in the majority of lignocellulosic waste, which also causes environmental pollution. However, the plant biomass residue considered as “waste” has the potential to be converted into many valuable products such as biofuels, various chemicals, cheap energy sources for fermentation, animal feeds and human nutrients [40]. Globally, 10.2% (50.3 EJ / year) of the total primary energy supply per year (TPES) provides from a wide range of biomass sources [41]. The application of biofuels originating from lignocellulose combines many positive factors. Supply security is at the top of these. Because the geographical distribution of the lignocellulosic raw material is more evenly distributed than fossil fuels.

Using lignocellulosic biomass as raw material, liquid fuels such as ethanol or pyrolysis oil, gaseous fuels such as biogas (methane), and electricity can be produced by thermochemical or biochemical processes.

Much of the research undertaken in this framework over the last two decades has involved efforts to increase the degradability of lignocellulosic biomass with the aim of efficiently converting cellulose to ethanol, methane, and hydrogen [42].

One of the most important aims of the lignocellulosic biomass conversion technologies is the fractionation of lignocellulose to cellulose, hemicellulose and, lignin, the three main components of its structure [43].

### 1.3.1. Structure of Lignocellulose

Lignocellulosic biomass, in general, contains 40-50% cellulose, 25-30% hemicellulose and 15-20% lignin, proteins, ash and some other extractable components [41, 42, 44].

The composition of these components may vary from species to species. These polymers are linked to each other at different degrees in a heteromatrix structure. This association varies depending on the type, species, and source of the biomass. The amounts of cellulose, hemicellulose, and lignin compared to the other components are the key factors that determining the optimum amount of energy conversion route of lignocellulosic biomass in different species. Lignocellulosic feedstocks have been subjected to rigorous pretreatment processes in order to be used as a hydrolyzable substrate by commercial cellulolytic enzymes or by enzyme-producing microorganisms, to convert into fermentable sugars [45].

Due to its chemical composition and mechanical properties, plant cell walls are a rich source of chemical and fermentable sugar for biofuel production, but these features make plant cell walls also important as biopolymer sources in many industrial applications [46].

Cellulose is the main component of the plant cell wall that provides structural support and is also found in bacteria, fungi, and algae. When present in the form of an unbranched homopolymer, cellulose is a polymer composed of  $\beta$ -D-glucopyranose linked via  $\beta$ - (1,4) glycosidic bonds. The degree of polymerization of the cellulose chains varies from 10,000 glucopyranose units in wood to 15,000 in wild-type cotton. These repeating units in the cellulose chain are cellobioses, which is a disaccharide as opposed to glucose in other glucan polymers. Cellulose chains are grouped into microfibrils that come together to form cellulose microfibrils. Long chain cellulose polymers are linked by hydrogen and vander Walls bonds, which allow the cellulose to be packaged as microfibrils. The microfibrils are covered by hemicellulose and lignin [45].

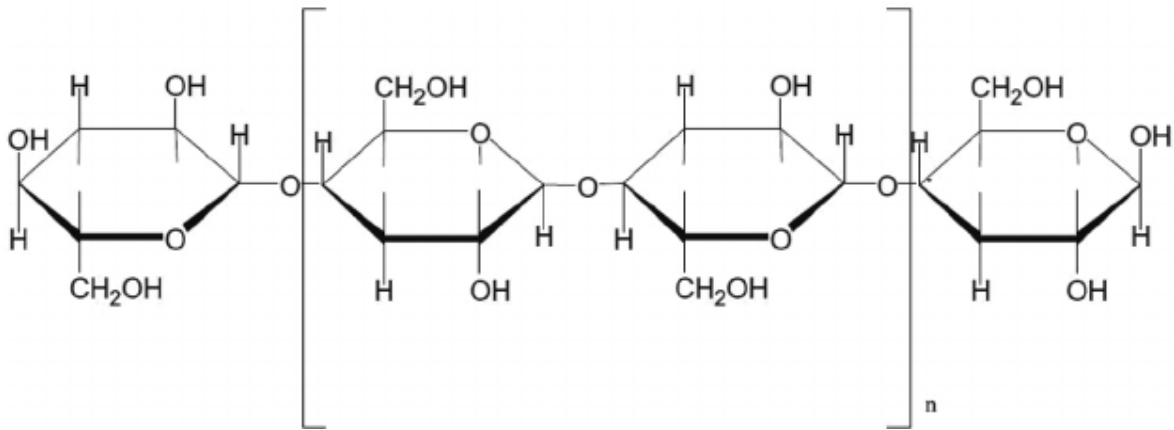


Figure 2. The chemical structure of cellulose [47].

Fermentable D-glucose can be produced from cellulose by breaking of  $\beta$ -1,4-glycosidic bonds with acid or enzyme activities. Cellulose in the biomass is found both in crystalline and amorphous form. Cellulose forms amorphous cellulosic cellulose in a low-level, undifferentiated cellulose chain, while the crystalline cellulose comes in large quantities. Cellulose in amorphous form is much more sensitive to enzyme degradation.

Cellulosic microfibrils are mostly independent, but ultrastructure of the cellulose is largely provided by covalent bonds, hydrogen bonds and, Van der Waals forces. The hydrogen bond in a cellulose microfibril determines the strength of the chain, but the hydrogen bonds between the chains can add stability (crystalline) or instability (amorphous) to the structure of the cellulose [45, 48].

Hemicellulose is the second most commonly found polymer, which accounts for 20-50% of the lignocellulosic biomass. It is not homogeneous like cellulose. Hemicellulose is branched into short chains of various sugars. These monosaccharide units contain xylose, rhamnose, arabinose, glucose, mannose, galactose and uronic acids. Hemicellulose backbone is a homo- or heteropolymer containing short branches, linked by  $\beta$ - (1,4) -glycosidic and rarely  $\beta$ - (1,3) -glycosidic bonds. In addition, hemicelluloses may be acetylated at a certain level such as in heteroxylan. It has a lower molecular weight when compared to cellulose [45, 49].

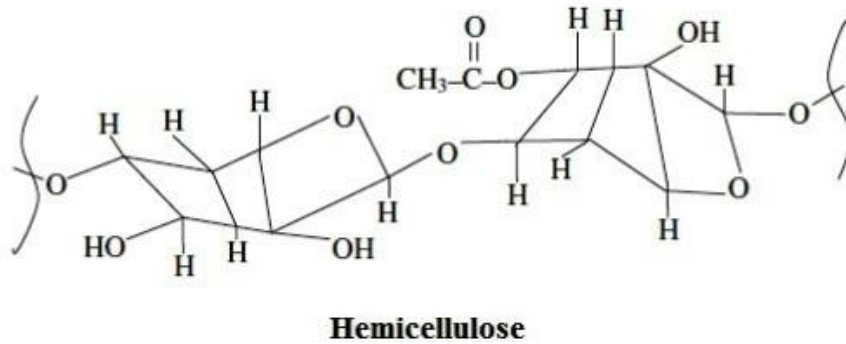


Figure 3. Structure of hemicellulose [50].

Lignin is the third most abundant polymer in nature. The complex consists of cross-linked phenolic compounds with a large molecular structure. The lignin in plant cell wall structure provides a rigid and impermeable resistance against stress and microbial attacks. Lignin contain three types of aromatic alcohols, coniferyl alcohol (guaiacyl propanol), Coumaryl alcohol (p-hydroxyphenyl propanol) and Sinapyl alcohol (syringyl alcohol). These monomers are linked together by Alkyl-aryl, alkyl-alkyl, and aryl-aryl linkages, respectively [45]. In addition, grass and dicot lignin contains phenolic acids such as p-cumaric and ferulic acid in large amounts which is esterified with alcohol groups [51].

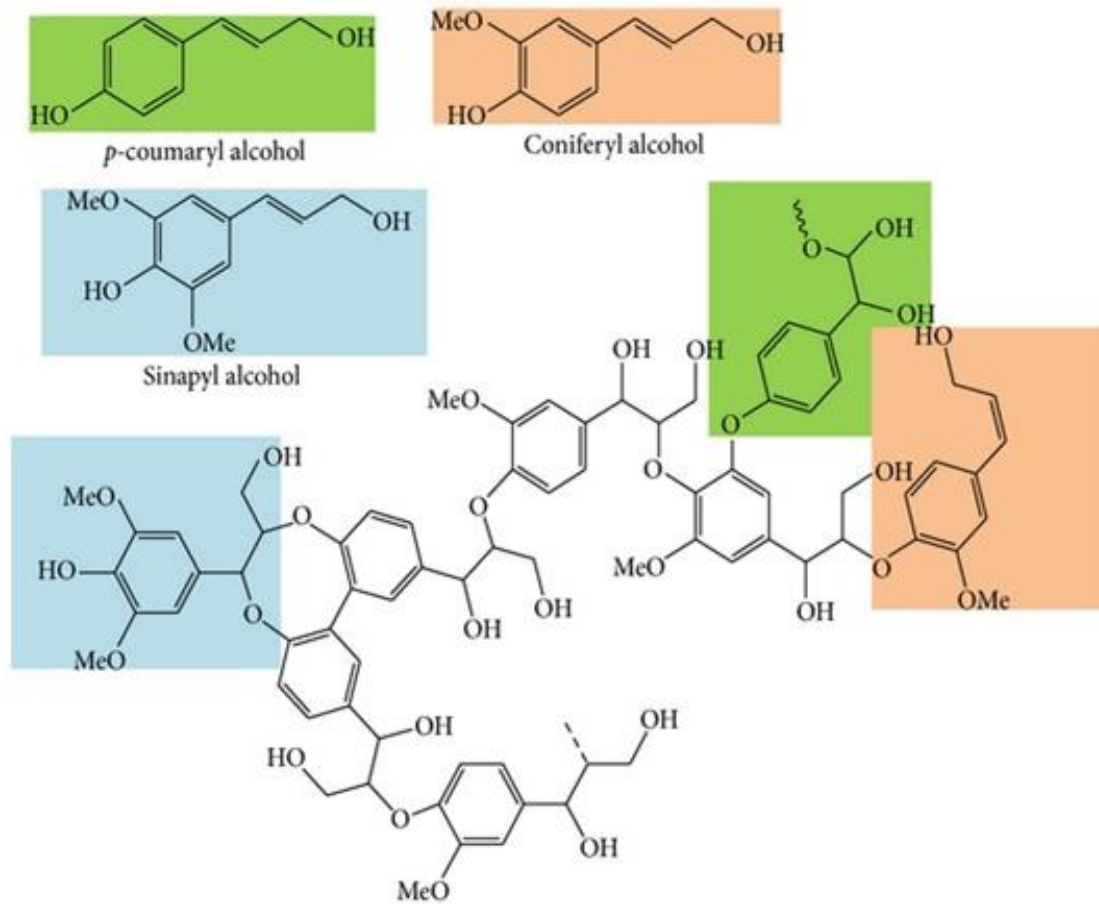


Figure 4. Representation of Lignin Structure [52].

While lignin is most resistant to the degradation, cellulose is more resistant to hydrolysis when compared with hemicellulose, because of its highly ordered crystalline structure [53].

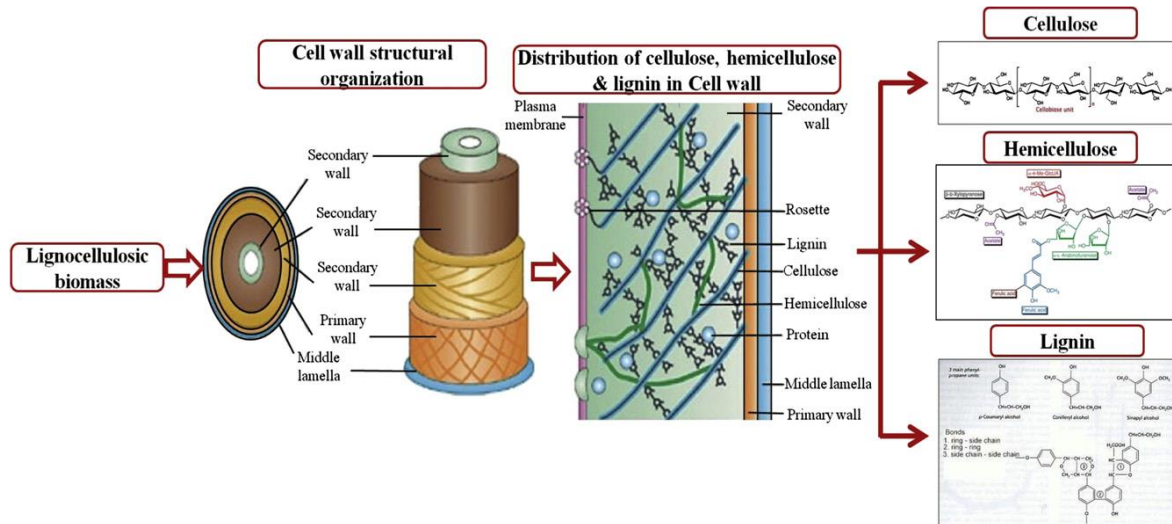


Figure 5. Schematic illustration of lignocellulose [42].

Today, the conversion of lignocellulosic biomass to valuable biochemicals such as ethanol still contains significant limitations.

Due to the complex structural and chemical properties of the lignocellulosic biomass, this resource is subjected to various processes in order to convert it into fuels and chemicals [38].

#### 1.4. Converting Lignocellulosic Biomass to Biofuels

Cellulosic components need to be converted to fermentable sugars in order to produce industrially valuable fuels and chemicals from lignocellulosic biomass [53]. Today, lignocellulosic biomass is processed by two main conversion pathways: biochemical and thermochemical process [54]. This can be achieved through three important technological processes; pyrolysis, gasification (thermochemical) and hydrolysis (biochemical) [38].

Gasification is a process that requires high temperatures ( $> 1100$  K). This high temperature is required for the formation of the endothermic synthesis gas [38, 55]. In this process, an oxidizing agent (oxygen, air, steam, etc.) is fed into the gasifier to partially burn the biomass. The obtained synthesis gas is subjected to different treatments, resulting in a variety of fuels including alkanes and methanol [56-60]. Although gasification can be used for the processing of all three components of lignocellulosic biomass, the low purity and highly diluted properties of the synthesis gas [61], negative effect of high moisture content

of the lignocellulosic biomass on the thermal efficiency of the process [62], and high biomass requirements [63] are among the limitations of this method.

Pyrolysis and liquefaction is the process of thermal decomposition of biomass at low temperatures (573-973 K) in an inert atmosphere [64, 65]. These processes convert the biomass into a black liquid known as bio-oil. This bio-oil is a complex mixture containing more than 300 highly oxygenated compounds, polymeric carbohydrates, and lignin fragments. Its water content is 25% [66, 67]. However, high oxygen content causes this material to contain low energy. In order to be used as transportation fuel, it needs to be deoxidized [68, 69]. Although pyrolysis is a simple process, the downstream processes required for deoxidization and the need to upgrading over the zeolite catalysts in order to produce aromatic fuels cause an increase in cost [70].

### **1.5. Ethanol Production From Lignocellulose**

The conversion of lignocellulosic biomass to fermentable sugars in large quantities is a potent application in the production of bioenergy [53, 54, 70].

Biomass composes from three main components; cellulose, hemicellulose, and lignin. The biochemical conversion process involves the hemicellulose fractionation process so that the cellulose can be more accessible during the reaction. During this process, the lignin remains unreacted. Lignin processing is accomplished by thermochemical conversion. Enzymatic hydrolysis (which also called saccharification) is an economical and effective method of obtaining fermentable sugars from the pre-treated lignocellulosic biomass [72-74].

The conversion of cellulosic biomass to ethanol via enzymatic hydrolysis involves the following steps; selection of suitable biomass materials and effective pre-treatment, biological conversion (production of saccharolytic enzymes such as cellulase and hemicellulase with auxiliary enzymes, hydrolysis and fermentation), post-processing (product recovery), public services and waste treatment [42]. Unlike starch-containing edible biomass, the complex structure of lignocellulosic biomass containing lignin-protected, high-grade crystallized cellulose constitutes a highly resistant impermeability shield that makes depolymerization via related enzymes and chemicals difficult [76, 77]. Consequently, pre-treatment refers to the processes of making cellulosic biomass suitable to the activity of hydrolytic enzymes [42]. The conversion of cellulose to ethanol basically varies on the

methods that used to achieve hydrolysis and fermentation. These two steps constitute the most important steps in ethanol production, although they involve newer and more advanced processing technologies than pre-processing technologies. The hydrolysis step can be classified into two categories according to the use of mineral acid and cellulase enzyme [75].

Acid and enzymatic hydrolysis are effective methods for separating lignocelluloses lignin and carbohydrate units into their monomeric components. These methods have the advantage that the biomass can be separated at lower temperatures than gasification and pyrolysis. Unlike starch-containing edible biomass, the complex structure of lignocellulosic biomass that containing lignin-protected, high-grade crystallized cellulose, constitutes a highly resistant impermeability shield that makes depolymerization via related enzymes and chemicals difficult [38].

Enzymatic hydrolysis is accomplished through a mixture of complex enzymes produced by microorganisms containing various bacteria and fungi. These microorganisms have the ability to reduce cellulosic biomass to glucose monomers. In bacteria, these complex enzyme systems, consisting of multiple subunits, form complexes called cellulosomes [53]. In the enzymatic hydrolysis of cellulose, cellulases; endoglucanases, exoglucanase or cellobiohydrolase and  $\beta$ -glucosidases act synergistically [78, 79]. In addition, glucuronidase, acetylsterase, xylanase,  $\beta$ -xylosidase, galactomannanase, and glucomannanase are also involved in the separation of cellulose into its monomeric units [79]. Among the bacteria; *Clostridium*, *Cellulomonas*, *Bacillus*, *Thermomonospora*, *Ruminococcus*, *Bacteroides*, *Erwinia*, *Acetovibrio*, *Microbispora* and *Streptomyces* [80], fungi; *Sclerotium rolfii*, *Phanerochaete chrysosporium*, *Trichoderma* species; *Aspergillus*, *Schizophyllum*, and *Penicillium* can produce cellulases [71, 79, 81].

Today, while acid-dependent processing technologies are more developed, enzymatic processing technologies are carried out with almost equal costs, and as related technologies are developed, they are expected to become more advantageous in terms of cost, in the future.

In addition, acid-dependent processing technologies have more negative environmental impact than enzymatic processing. The only significant advantage of the acid-dependent process relative to the enzymatic process is that acid dependent hydrolysis equally effective in soft and hard wood since enzymatic hydrolysis is readily applicable to hardwood and herbaceous materials when compared to softwood [75]. Due to certain structural differences between hardwood and softwood, hardwood has a less recalcitrance to enzymes than softwood [82].



Fermentation takes place after enzymatic hydrolysis of the biomass with the corresponding enzymes. Following hydrolysis of biomass to fermentable sugars, these sugars are converted to ethanol by several microorganisms.

In order to produce commercially valuable ethanol, an ideal microorganism must possess certain properties. The microorganism to be used in ethanol production is desired to have the ability to use a wide variety of different substrates, to have the ability to produce ethanol with high yield, titer and productivity, and should harbor high tolerance to ethanol, temperature and to inhibitors in the hydrolyzate. Today, the most used microorganisms in ethanol production are *S. cerevisiae* and *Z. mobilis*. *S. cerevisiae* has been used commercially to produce ethanol from starch and sucrose but this yeast is only capable of producing ethanol from hexose sugars [74].

Cellulase production of microorganisms is controlled by some genetic mechanisms. In this context, various studies have been carried out including the development of strains through mutations in order to increase the production of cellulases. However, many of these studies remain with limited success. Thus the biological processing technology is still a promising option but still presents significant challenges [42]. The elucidation of the molecular mechanisms leading to lignocellulosic biodegradation and the enhancement of bioprocessing potentials of cellulolytic microorganisms can be achieved through genetic engineering based approaches [53].

Among these, CBP has an outstanding potential as an alternative approach, which has been offering the ability to perform cellulose hydrolysis and fermentation in one step [71].

## **1.6. Consolidated Bioprocessing (CBP)**

Biologically based technologies in energy conversion, especially the conversion of lignocellulosic biomass into biofuels, are promising. Although conversion technology, including a step toward cellulase enzyme production, has become the focal point of many investigators, the CBP method, which offers cellulase production, cellulose hydrolysis and fermentation in a single step, has an extraordinary potential as an alternative approach [4, 83]

Cellulase production, hydrolysis, and fermentation of both cellulose and hemicellulose hydrolysis products are performed in a single step in the consolidated bioprocessing. Consolidated bioprocessing is defined as the processing of lignocellulosic biomass in a

single step, followed by cellulase production, hydrolysis, and fermentation, following the pre-treatment step. The term “consolidated bioprocessing” was first proposed by Lynd in 1996 [84] as an alternative to “direct microbial conversion” (DMC) [85] and entered into the literature [7].

As mentioned earlier, bioconversion of cellulosic biomass by enzymatic hydrolysis occurs in four basic steps; first cellulase is produced, then following the cellulose hydrolysis, the soluble cellulose and its hydrolysis products and the soluble hemicellulose and its hydrolysis products are fermented.

Depending on which of these steps are consolidated, alternative processing arrangements can be made. Separate hydrolysis and fermentation (SHF) is a process that involves four distinct steps and four different bio-catalyzer. Simultaneous saccharification and fermentation (SSF) combine the hydrolysis of the cellulose and fermentation of its hydrolysis products in a single step, in addition to this step, two further steps are carried out, namely, (1) production of cellulase and (2) fermentation of hemicellulose hydrolysis products.

Simultaneous saccharification and co-fermentation (SSCF) take place in two steps; cellulase production and cellulolytic hydrolysis and formation of hydrolysis products [7].

Lynd et al [7] defined a cost-benefit view, considering the same amount of substrate (a fraction of feedstock used to produce the enzyme required for fermentation) used for cellulase production,  $F$ , while the remaining feedstock,  $F-1$ , used to produce the desired products. Accordingly,  $F$  values in SSF and SSCF-focused processes range from 0.03 to 0.05 (the amount of enzyme equivalent to 10-15 FPU / g cellulose hydrolysis). However, for aerobic cellulase production, a minimum average of 300 FPU/g of cellulose is used [86] and production occurs at the end of a 5-7 day reaction period [87, 88]. In this context, the optimum  $F$  value in terms of economical view can only be achieved with the minimizing the total cost of the cellulase production and other costs associated with cellulase production, which in turn results loss of yield [7, 89].

Cellulase production in SSF, SCF, and SSCF is achieved with an aerobic microorganism due to the ability to produce higher yields of cellulase due to higher ATP production. However, the conversion of hydrolysis products into fermentable monomers for use as fuel is carried out by anaerobic microorganisms. In this respect, avoiding of cellulase production cost is one of the most important advantages of CBP process [1, 7, 90].

CBP is a potential method to find an application field that allows any desired fermentation product to be obtained at low cost through the processing of cellulosic biomass, but in order to use this method efficiently, appropriate organisms need to be modified in this direction. It is expected that an organism that can be used for this purpose under industrial conditions will have the ability to use both the substrate effectively and produce the desired product at high yield and titer [83].

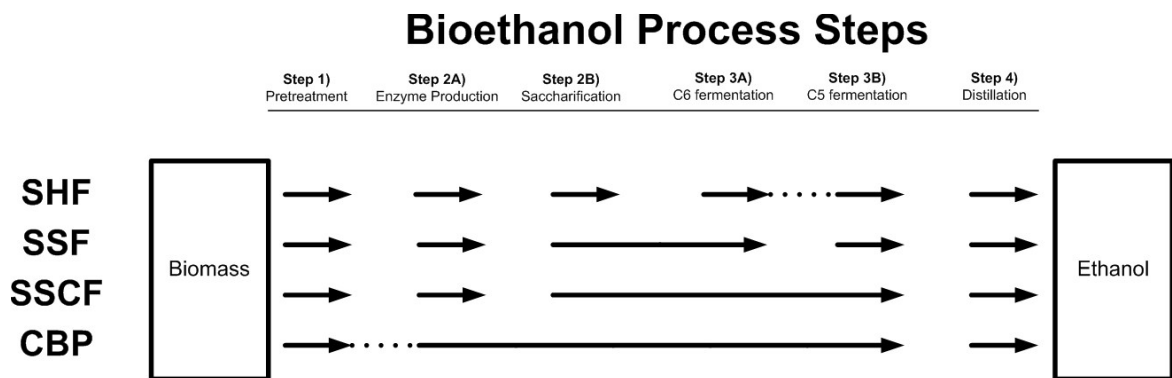


Figure 6. Bioethanol production process steps, SHF: separate hydrolysis and fermentation; SSF: simultaneous saccharification and fermentation; SSCF: simultaneous saccharification and co-fermentation; CBP: consolidated bioprocessing; C6: hexose; C5: pentose [3].

Since a native organism with these properties is not found in nature, genetic and metabolic engineering approaches and microorganisms with these properties need to be developed. With the aim of developing microorganisms suitable for industrial use through the means of genetic engineering, two strategies come to the forefront: (1) native strategy (starting with microorganisms that are naturally capable of converting cellulosic biomass into soluble monomers) (2) recombinant strategy (starting with microorganisms that do not naturally capable to utilize the components of lignocellulosic biomass and thus need modification with genetic engineering). Both strategies are finding application areas for obtaining different products [4]. Although the CBP method can in principle be applied to produce a wide range of products from plant biomass, CBP has drawn on the greatest interesting natural strategy to naturally produce microorganisms capable of lignocellulosic disintegration on industrial conditions and for the first time commercially applied to ethanol production [6].

In the low-carbon future, there is a wide consensus that biomass will play an important role [13], and therefore the production of transport fuels is a priority for the use of biomass [12]. For this purpose, the main problems encountered in the natural strategy are; the inadequacy of tools to perform the desired genetic modifications and the inability of these tools to be applied to the desired microorganism in a manner that is robust to tough industrial processes, yielding products at high yield and concentration [6]. Especially in the last decade, improvement of microorganisms that can be able to converting hemicellulose-based sugars into ethanol with high yields were noted as the most important progression both in biomass conversion and metabolic engineering.

In this course, *Escherichia coli* and *Klebsiella oxytoca* [91, 92], which have the ability to produce high ethanol, can be exemplified as microorganisms developed by the native strategy. These organisms can natively use hemicellulose-based sugars, however they were engineered to produce ethanol with high yields. Similarly, *Saccharomyces cerevisiae* [93, 94] and *Zymomonas mobilis* [95, 96] can be shown among the microorganisms developed as an example of a recombinant strategy [83].

Lignocellulosic biomass has a number of properties that must be considered in an ethanologic organism for commercially successful ethanol production. These; (1) minimal by product formation (2) high yield and productivity ( $> 1 \text{ g / L / h}$ ) (3) safe (4) wide tolerance to environmental conditions (5) high ethanol tolerance (6) (12) producing low biomass (13) with low nutrient requirements (8) high cellulolytic activity (9) substance tolerant (10) high solids feeding and substrate loading tolerance (11) easy genetic modification. Besides, some features expected in the processing process are; (1) $> 90\%$  theoretical yield (2) higher ethanol concentration ( $> 5\%$  (v / v) (3) the minimum number of processing steps (4) renewable cells (5) common fermentation of substrates (6) limited pre- (7) unrestricted treatment or untreated treatment. Today, a single organism that contains all these features unfortunately does not exist. With *S. cerevisiae* and *Z. mobilis* being strong ethanol producers, the substrate variety is very limited. *Z. mobilis* also has a complex pathway of destruction of unwanted end products during fructose and sucrose destruction [3].

Studies using classical mutation techniques for obtaining stable strains capable of producing high-yield ethanol have not been successful [7, 16]. The need for genetic modifications to enable *S. cerevisiae* and *Z. mobilis* to use in the production of second generation fuels has attracted attention to the thermophilic microorganisms associated with the oil crisis of the 1980s [3]. Thermophilic bacteria have many desirable properties for the

second-generation fuel production listed above [8, 97]. Thermophiles have a much broader range of carbohydrate breakdown characteristics compared to *S. cerevisiae* and *Z. mobilis* and do not require extensive rinsing, cooling and heating for cultivation in the fermentation tank. It is also possible to recover ethanol directly from the fermentation liquid by *in situ* vacuum distillation. Beyond the ability to run at different temperatures in a very wide range of thermophiles, they can tolerate low pH and high salt concentrations during fermentation with low nutrient requirements [1]. In addition, all known thermophiles are safe microorganisms in the lowest microbial risk class. When viewed from the point of view of processing, high temperature processing also facilitates the mixing process because it reduces the viscosity and increases the substrate loading rate. Mass transfer rates at higher temperatures are higher and the risk of mesophilic contamination is lower [3].

### **1.7. Metabolic Engineering of Thermophilic Bacteria For Ethanol Production**

There is a wide consensus that biomass will play an important role in the supply of energy [13] and transport fuels [12]. Lignocellulosic biomass is the most commonly available renewable energy source in nature, but the number of microorganisms that have the ability to use this material as a carbon and energy source is few [83]. Among them, thermophiles are the most interesting microorganisms for the application of CBP due to the direct proportion between the hydrolysis of crystalline cellulose and the temperature [98]. Today, the vast majority of bioethanol production is accomplished through mesophilic organisms. However, thermophilic bacteria are more advantageous than mesophiles in terms of eliminating certain troubles such as the possible inhibitor effects of sugar derivatives being avoided and high ethanol titre being obtained [3]. As mentioned above, there are two strategies for the development of the microorganisms for CBP; native and recombinant strategies. The microorganisms to which the native strategy can be applied could be fall into three categories; fungi, free-enzyme bacteria and cellulosome-forming bacteria [6].

Anaerobic thermophiles are at the forefront of microorganisms in which the natural strategy finds application [4]. The reasons for this, include a complex enzyme system known as a "cellulosome" that breaks down a variety of substrates such as cellulose, starch, xylen, and even lignin, all of which are naturally present in the lignocellulosic biomass [6, 7]. Complex enzyme systems exhibit synergistic activity on lignocellulosic biomass that is much more effective and efficient than an enzyme alone can show [7]. Among these

anaerobic bacteria identified as ethanologen, there are mainly *Clostridium* (in particular; *Clostridium thermocellum*), *Caldanaerobacterium*, *Thermoanaerobacter* (in particular; *Thermoanaerobacter mathranii*) and *Thermoanaerobacterium* (in particular; *Thermoanaerobacterium sachharolyticum*, *Thermoanaerobacterium xylanolyticum*, *Thermoanaerobacterium thermosaccharolyticum*) genus members [3].

As a result of studies with *Clostridium thermocellum* belonging to *Clostridium* genus and *Thermoanaerobacterium saccharolyticum* belonging to genus *Thermoanaerobacterium*, many techniques for applying gene transfer systems and genetic engineering in these bacteria have been developed and successfully applied [15, 17, 18]. Cellulolytic thermophiles such as *Clostridium thermocellum* and hemicellulolytic thermophiles belonging to the genus *Thermoanaerobacter* and *Thermoanaerobacterium* can be used together for the hydrolysis and fermentation of the biomass sugars [4]. The use of thermophilic hemicellulosic microorganisms may reduce the amount of cellulase needed for hydrolysis [16, 99]. In addition, processing at high temperatures could reduce the degree of heat exchange and could reduce the risk of contamination, both after pretreatment and prior to distillation [7]. However, most of the organisms belonging to this class, naturally, can not achieve a homogeneous ethanol fermentation and do not have high product tolerance.

Comprehensive efforts have not been successful using classical mutagenesis techniques to obtain stable strains exhibiting high ethanol yield [16]. Metabolic engineering based strain development studies are needed to address these deficiencies. The realization of these studies requires research into the elucidation of the metabolism of thermophilic microorganisms [15]. These microorganisms are able to grow between 45-65 °C and pH 4.0-6.0. The use of anaerobic thermophiles in industrial biotechnology offers potential advantages over mesophilic organisms, such as lower risk of contamination, higher reaction rate and reduced heating and cooling costs. [27]. Thus genetic and metabolic engineering of such thermophiles has a great potential to produce important fuels from biomass via CBP.

### **1.7.1. Metabolic Engineering of *T. saccharolyticum* for Ethanol Production**

*T. saccharolyticum* is an anaerobic thermophilic microorganism with the ability to use sugars such as cellobiose, glucose, xylose, mannose, and galactose and arabinose present in the cellulolytic biomass and capable of hydrolyzing xylan, mannan, starch and pectin but not cellulose, and produces ethanol, acetate, lactic acid, CO<sub>2</sub>, and H<sub>2</sub> as fermentation products

[28]. These features make this bacterium a remarkable model organism in that genetic manipulations can be performed [23, 100]. Previously, these bacteria have been modified to produce high yields of ethanol with deletion of genes that are effective in the production of organic acids and H<sub>2</sub> [16, 30]. *T. saccharolyticum* JW/SLYS485 is a gram-positive bacterium, was isolated from Yellowstone National Park, Wyoming [101], and designated as *Thermoanaerobacterium saccharolyticum* JW/SL-YS485 [102]. Genetic modification of *T. saccharolyticum* by metabolic engineering has been initiated with the blocking of lactate production by deletion of lactate dehydrogenase gene, which was resulted in 5% increase in the yield of ethanol production, and immediately followed by the disruption of the phosphotransacetylase (pTA) and acetate kinase (ack) (acetate) genes, resulted in strain ALK1 which can produce ethanol at the theoretical maximum level of 90-100% [16, 28]. After continuous cultivation of this strain (ALK1) for up to 3000 hours, strain ALK2, which reached of 92% ethanol yield, a titer of 33 g/L and with 2.2 g/L/h productivity, was obtained [16]. Since both genetic modifications in the ALK2 strain are carried out by chromosomal integration of an antibiotic resistance gene and until now only two antibiotic markers (kanamycin and erythromycin) have been available in *T. saccharolyticum*, no further genetic modification could be performed with this strain. This problem has been overcome by reconstructing the strain carrying the deletion of *ldh* and *pta-ack* with a marker recycling strategy [103]. The resulted new mutant strain M0355 showed similar ethanol yields (~ 94% yield, 25 g/L titer, and 1.13 g/L/h productivity) with ALK2, but salt accumulation in this new strain has been found to limit the growth of *T. saccharolyticum*, posing a problem in controlling pH [104]. While this strain was not produced organic acids in excessive quantities, the uptake of ammonium was led to the acidification of the medium. To overcome the need for pH control, M0355 was modified to use urea as a source of nitrogen [105]. After a series of genetic modifications carried out on this strain, strain M1442 was obtained with 90% of theoretical yield, 61 g/L titer and 2.13 g/L/h productivity [106].

Most of the anaerobic bacteria produce organic acids, H<sub>2</sub> and CO<sub>2</sub> as well as ethanol by mixed fermentation. Theoretically it is possible to produce two moles of ethanol from per mole C<sub>6</sub> sugar (1) (C<sub>6</sub>H<sub>12</sub>O<sub>6</sub> → 2 C<sub>2</sub>H<sub>6</sub>O + 2 CO<sub>2</sub>) [107]. The mechanism of the distribution of the products in mixed acid fermentation is still unknown but it is hypothesized that it is determined by the mass action principle. Studies suggesting that the increase in partial pressure of H<sub>2</sub> leads to an increase in ethanol production [108, 109] or that a decrease in

ethanol production is followed by a decrease in H<sub>2</sub> partial pressure seems to support this hypothesis [110].

Another hypothesis is that some specific regulatory processes determining product distribution. An example of this assumption is the Rex system, which was first discovered in *Streptomyces coelicolor* [111]. Rex protein regulates gene expression depending on the ratio of NADH / NAD<sup>+</sup> and it was shown that deletion of the *rex* gene increased ethanol production in *Staphylococcus aureus* and *Clostridium acetobutlicum* [112-114]. In another study conducted in *Clostridium kluyveri*, it was determined that Rex protein controls the synthesis of butyrate and electron bifurcation [115]. In *Thermoanaerobacter ethanolicus*, a redox-sensing protein (RSP) has been reported to repress the transcription of alcohol dehydrogenase [116].

In *T. saccharolyticum* as a result of pyruvate metabolism, a reduced ferredoxin is produced and the ferredoxin could be reoxidized by hydrogenase to produce molecular hydrogen or these electrons could be used for ethanol production because in theoretical, a decrease in hydrogen production should result in an equivalent increase in ethanol production [6]. Another study that was carried by Zheng et al. has been suggested that a hydrogen sensing *hydS* gene which is homologous of the *hfsB*, transduce the signal by its PAS domain to a Ser/Thr protein kinase present in the hydrogenase operon and thus the redox potential is balanced via transcriptional regulation. PAS domains are mostly present in organisms and functions as redox, light, and oxygen sensing apparatus [117].

Recently this strategy (called electron centered metabolic engineering) was applied to the engineering of the wild-type *T.saccharolyticum*'s *hfs* hydrogenase [29, 30]. *T. saccharolyticum* metabolism consists of three putative hydrogenases; a NAD-dependent [FeFe]-hydrogenase (*hyd*), a probable [Ni-Fe] hydrogenase (*ech*) and a Fe-only-hydrogenase which are assigned as *hfs* (hydrogenase-Fe-S) containing Fe-S clusters and a PAS (*Per-Arnt-Sim*) domain. In the study which was conducted by Saw et. al. [30] the whole *hfs* operon was disrupted with a *kan* marker replacement, and it was seen that removal of only the *hfsABCD* was not sufficient to produce higher ethanol, even the acetic acid reduced, and 95% less hydrogen produced in the  $\Delta hfsABCD::kan$  mutant (HK03) when compared to the wild-type strain. Also, a growth deficiency was observed in the mutant strain, probably caused by the difficulty of balancing the redox potential in response to the imbalanced end-product formation and the strain was produced almost 1.6 fold more lactate than the wild-type. A 44% ethanol yield increase was obtained only after deletion of the *ldh* gene in the



$\Delta hfsABCD::kan$  strain. Later the same group was found that introduction of two point mutations which was in *hfsB* (was caused to a frameshift) and in *hfsD* (was caused by an amino acid substitution) subunits, into the *hfs* operon, resulted in a 50% increase in ethanol production [29].

In order to better understand the mechanism of *hfs* hydrogenase, the functions of the genes and their effects on ethanol production need to be elucidated. In this study, it is aimed to understand the role of the *hfs* hydrogenase in the wild-type phenotype of *T. saccharolyticum* via deletion of the each of gene in the *hfsABCD* operon.

### 1.7.2. Metabolic Engineering of *C. thermocellum* for Ethanol Production

*C. thermocellum*, one of the most anaerobic bacteria that form cellulose, is one of the most developed places for CBP applications. This is due to the ability of *C. thermocellum* to destroy high cellulose. Recent studies on the metabolic properties of *C. thermocellum* [17] and advances in techniques developed for genetic modification [18, 118, 119] have shown that *C. thermocellum*, as a CBP organism, has become an attractive way to obtain biofuels or biochemical products [21]. *C. thermocellum*, an anaerobic bacterium, is capable of degrading cellulosic debris to an adequate extent, especially through the complex enzyme system called "cellulosome" attached to the cell wall [120, 121]. Cellulosomal enzymes break down the cellulose into cellodextrin, which can be converted to ethanol, acetate, L-lactate, formate, hydrogen gas and CO<sub>2</sub> by acid fermentation [15, 16]. However, wild-type *C. thermocellum* is an organism that produces ethanol in low yield and concentration with acetates, lactates, H<sub>2</sub>, formate and various amino acids as by-products [122]. In addition, ethanol tolerance of wild strains is low [123]. Recent methods developed for genetic modification of *C. thermocellum* [9, 119, 124] have increased the likelihood that this bacterium can produce a variety of industrial fuels from cellulosic substrates. Although the application of these methods still has some difficulties, it can now be successfully used to develop industrially valuable features in *C. thermocellum*. Theoretically maximum ethanol yield; is defined as an ethanol per pyruvate. In organisms, pyruvate ethanol production can be accomplished by means of four different metabolic pathways (named after considering the key enzymes that are catabolizing pyruvate); (1) pyruvate decarboxylase (PDC), (2) pyruvate dehydrogenase (PDH), (3) pyruvate formate lyase (PFL) and (4) pyruvate ferredoxin dehydrogenase (PFOR). The PDC pathway is a pathway in mesophilic organisms

[30, 31] and is rarely seen in thermophiles. The PFL pathway is the pathway used for the purpose of ethanol production in *E. coli* [125]. The PDH path is available in *Geobacillus thermoglucosidasius* [126]. PFOR is the path to which many obligatory anaerobes, including *C. thermocellum*, are used. All known anaerobic saccharolytic thermophiles, including *T. saccharolyticum* [27], produce organic acid as well as fermentation ethanol. The lactic acid is formed by L-lactate dehydrogenase, while acetic acid is formed by pyruvate, pyruvate-ferredoxin oxidoreductase (PFOR), phosphate acetyl transferase and acetate kinase (Figure 7) [7, 16]. It is assumed that the theoretical maximum ethanol yield for all organisms defined in this study is one ethanol per pyruvate, while NADH is derived from glyceraldehyde-3-phosphate dehydrogenase (GAPDH) in the glycolysis reaction [15].

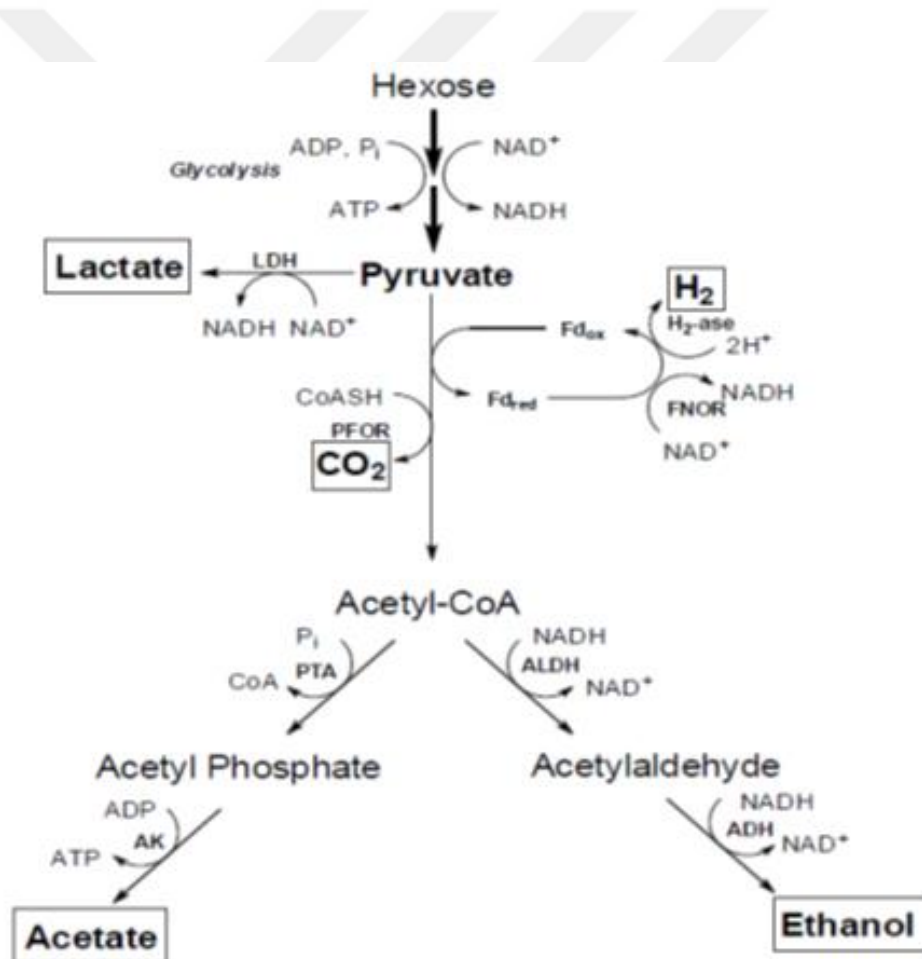


Figure 7. A schematic representation of the products formed after pyruvate in glucose degradation in obligative anaerobes, ALDH: acetaldehyde dehydrogenase; ADH: alcohol dehydrogenase; AK: acetate kinase; FNOR: ferredoxin oxidoreductase; H<sub>2</sub>-ase: hydrogenase; LDH: lactate dehydrogenase; PFOR: pyruvate ferredoxin oxidoreductase; PTA: phosphotransacetylase [3].

Reduced ferredoxine is formed during the conversion of the pyruvate via PFOR to acetyl-CoA (Figure 8.). The electrons from Ferredoxine can be used in the production of H<sub>2</sub> as well as in the production of ethanol, and it is theoretically expected that a decrease in hydrogen production would cause an equivalent increase in ethanol production. Along with being the main soluble products competing with acetic acid and lactic acid H<sub>2</sub> production is a major cause of ethanol production, as it is the leading pathway for electron transfer.

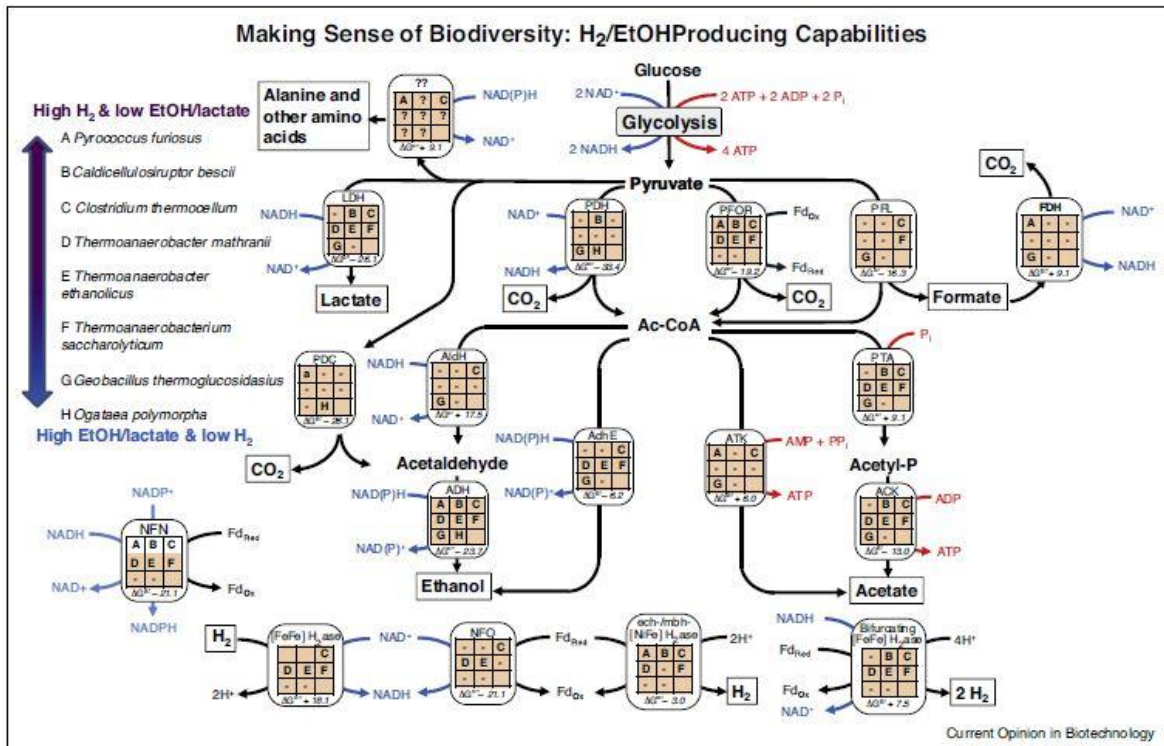


Figure 8. Schematic illustration of ethanol production pathway in thermophiles [15].

The use of protons as electron transporters instead of glycolytic intermediates led to the production of acetyl-CoA, acetate and ATP. This, in turn, leads to a decrease in the electrons required for the reduction of acetyl-CoA, while also meaning more energy for the cell than it can use. H<sub>2</sub> production in the cell is catalyzed by an enzyme class called "hydrogenases". Hydrogenases are separated by three groups according to the metal ion they contain in active sites; [Fe] hydrogenases, [FeFe] hydrogenases and [NiFe] hydrogenases [100], which have been detected only in methanogens until now. *C. thermoceillum* contains three predicted [FeFe] hydrogenase and a ferredoxine-dependent [NiFe] hydrogenase (*ech*) [127].

With the aim of reducing the production of hydrogen in *C. thermocellum* and transferring the ethanol production to the theoretical levels. In a study conducted by Biswas et al. [128], the *hydG* gene, which catalyzes the conversion of the holoenzyme of the [Fe-Fe] hydrogenase apoenzyme to the four genes responsible for H<sub>2</sub> production in *C. thermocellum*, was inactivated by deletion of the *ech* and the hydrogenase activity was completely eliminated, the production was increased to 64% of the theoretical level (yielding 90% higher ethanol from wild strains). However, it has been found that the reason for the imbalance of hydrogen metabolism to occur in the redox potential is that the new strain ethanol tolerance is the same as the wild type and shows some growth deficiencies. Product tolerance is one of the key factors in the implementation of KBP. Recently there have been studies to increase the ethanol tolerance of *C. thermocellum* [129]. In this study, it was determined that the selected strains were able to grow in ethanol concentrations exceeding 60 g / L. This concentration is sufficient to avoid disadvantage compared to other producers traditionally used in ethanol production of a thermophilic organism, given the lignocellulosic conversion [75]. However, both *C. thermocellum* and other thermophiles produce ethanol at a concentration of  $\leq 26$  g / L [7], suggesting that this barrier needs to be overcome by metabolic engineering methods to achieve maximum ethanol production [4].

In the present work it is aimed to restore the redox balance in *C.therhermocellum*  $\Delta hydG\Delta ech$  strain (LL1147) by heterologues expression of the mutated *hfs\** operon from *T. saccharolyticum* (strain LL1189) under three different promoter to obtain a high yield ethanol producing ethanologen strain for industrial purposes.

### 1.8. Fe-Only [FeFe] Hydrogenases

Hydrogenase is the name given to the class of metalloenzymes that catalyze the oxidation and production of molecular hydrogen (H<sub>2</sub>).

It was first discovered in 1931 by Stephenson and Stickland [133] in the cell free extracts of some bacteria that can catalyze the reversible oxidation of hydrogen in the presence of an electron carrier such as methyl viologen and named as hydrogenase.

The direction of this reaction (2) ( $2H^+ + 2e^- \leftrightarrow H_2$ ) depends on the redox potential of the components which the enzyme is active. If an electron acceptor is present, the direction of the reaction will proceed in the direction of H<sub>2</sub> oxidation, whereas if an electron donor is present, the reaction will proceed in the direction of H<sub>2</sub> release [133]. Molecular hydrogen

is a metabolite that involves a series of reaction in wide range of microorganisms [134]. This enzyme, which is commonly present in anaerobic microorganisms, functions to remove excess electrons during fermentation or to electron uptake for reduction reactions in the course of energy production [135].

Hydrogenases (EC 1.18.99.1), are classified according to the structure of their catalytic sites [136], [Fe] hydrogenases, [FeFe] (Fe-only) hydrogenases and [NiFe] hydrogenases [100]. Their common feature is the presence of a central Fe atom coordinated by a ligand. Here, the Fe-atom, which is catalytically active, represents the diiron-moiety which is covalently bonded to the [4Fe 4 S] cluster. The combination of [FeS] and this diiron cluster is called the H-cluster, and this combination leads to a very high efficiency rate of [FeFe] hydrogenases. While [NiFe]-hydrogenases can be found in all major prokaryota, [FeFe]-hydrogenase mainly can be found in the classes of *Clostridia*, *Proteobacteria*, and *Thermotogae* [137].

## **1.9. Biofuel in Turkey**

Almost all countries are intensively engaged in scientific and R&D studies aimed at bringing alternative sources of energy into use in the biofuels market, particularly in the USA and Brazil, and in the European Union, headed by Germany [138, 139].

Primary generations of biofuel production are being carried out in Turkey together with developing technologies since 2000's. Turkey's first commercial biofuel application began in 2005 [140]. The vast majority of raw material of the produced biofuels are imported [141]. Turkey's forest-based biomass potential, total amount of forest-derived waste is 4 million 800 thousand tons [139]. Biofuel is mostly produced as biodiesel.

## 2. MATERIALS AND METHODS

### 2.1. Chemicals and Kits

All the chemicals used in this study were molecular biology grade and were prepared with distilled water (Millipore, Billerica, MA, USA).

Phusion® High-Fidelity PCR Master Mix with HF Buffer (M0531L) and NEBuilder® HiFi DNA Assembly Master Mix (E2621L) were provided from New England Biolabs.

RT-PCR was performed by iScript cDNA synthesis kit (170-8891), SsoFast™ EvaGreen® Supermix (172-5201) from Bio-Rad.

For RNA experiments, Qiagen's RNeasy Mini Kit (74104), QIAprep Spin Miniprep Kit (27106), RNaprotect Bacteria Reagent (76506) and RNase-Free DNase Set (79254) and Proteinase K (19131) were used.

Agar (400400010), LB Broth, Miller (611875000) and Nickel (II) chloride hexahydrate (AC207675000) were supplemented from Acros Organics.

Other chemicals including Ammonium chloride (213330), Ammonium sulfate (A5132), Yeast extract (Y1625), Calcium chloride dihydrate (C5080), Citric acid monohydrate (C0706), D-(+)-Cellobiose (C7252), D-(+)-Xylose (X1500), DL-Dithiothreitol (43816), Sodium citrate tribasic dihydrate (S4641), L-Cysteine hydrochloride monohydrate (C6852), Magnesium chloride hexahydrate (63065), Manganese chloride (M3634), MOPS Sodium salt (M9381), Resazurin sodium salt (R7017), Sodium molybdate dihydrate (M1003), Sodium Sulfate (2391313), Sodium bicarbonate (S5761), Cobalt (II) chloride (C8661), Cooper (II) chloride hydrate (C3279), Iron (II) chloride tetra-hydrate (44939), Acetaldehyde (Sigma 402788), Acetyl coenzyme A sodium salt (A2056), Benzyl viologen dichloride (271845),  $\beta$ -nicotinamide adenine dinucleotide reduced disodium salt hydrate (N8129),  $\beta$ -Nicotinamide adenine dinucleotide 2'-phosphate reduced tetrasodium salt hydrate (N7505), Erythromycin (E5389), Biotin (B4639), Pyrodoxoxine HCl (P9755), Thiamine hydrochloride (T1270), Vitamin B12 (V2876), 4-Aminobenzoic acid (429767), 2-Floro-2'-deoxyuridine (F0503), Chloroform (Sigma C-2432), Lysozyme (L-6876), Isoamyl alcohol (I-9392), Phenol (P-4557) were provided from Sigma.

LB Agar, Miller (BP14252), Boric acid (BP168500), Potassium Citrate Monohydrate (P218212), Potassium Phosphate monobasic (P380212), Tris base (BP153-500), Urea

(U15500), Maganese (II) chloride tetrahydrate (M87100), Zinc (II) chloride (Z33100), Coomassie Protein Assay Reagent (23200) were purchased from Fisher Scientific

Some other reagents used in this study were; Carbenicillin disodium salt (Invitrogen 10177012), Kanamycin sulfate (Gibco 11815024), 8-azahypoxanthine (TCI A0555), TE buffer (Ambion 9858), , DNase-free RNase I (Epicentre N6901K) and DNA Clean & Concentrator™ (Zymo Research).

For GC assays 10% Carbon Dioxide (CO<sub>2</sub>) (1013, Nitrogen as balance) and 99,99 % Hydrogen (H<sub>2</sub>) (1049, methane as T.H.C) were used from Mesagas

## 2.2. Buffers and Stock Solutions

Thiamphenicol (Tm): 50 mg TM dissolved in 5 mL dimethyl sulfoxide (DMSO), filter sterilized (10 mg/ml) and 6 µg/mL used in colony selections.

5-Fluoro-2-deoxyuridine (FUDR): 50 mg FUDR dissolved in 5 mL dH<sub>2</sub>O and filter sterilized. 10 µg/mL was used in colony selection at a final concentration.

8-Azahypoxanthine (8AZH): 100 mg 8AZH dissolved in 1 M NaOH and filter sterilized. 500 µg/mL was used in colony selection.

CTAB/NaCl (hexadecyltrimethyl ammonium bromide): 4.1 g NaCl was dissolved in 80 mL of water. 10 g CTAB was slowly added while heating (≈65°C). and stirring. Final volume was adjusted to 100 mL and sterilized by autoclave.

Vitamin Solution (1000×): 50 mg Pyridoxamine HCl, 5 mg biotin, 10 mg p-amino benzoic acid, 5 mg Vitamin B12 dissolved in 25 mL water.

## 2.3. Strains, Media Composition and Growth Conditions

Strains and vectors used in this work and their sources were given in Table 1.

Table 1. Plasmids and strains used in this study

Strain and Plasmid ID	Description	Parental Strain	Accession number	Reference
LL1025	wt	<i>Thermoanaerobacterium saccharolyticum</i> JW/SL-YS485	SRA234880	[142]
LL1049 (M1442)	$\Delta(\text{pta-ack}) \Delta\text{ldh}$ $\Delta\text{or795}::\text{metE-ure} \Delta\text{eps}$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP052455	[16]

Continuous of Table 1.				
LL1187	$\Delta hfsABCD::kan$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP085668	[30]
LL1189	$hfsABCD::hfsABCD*-kan$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP085677	[29]
LL1267	$\Delta hfsA::kan$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP108081	this study
LL1268	$\Delta hfsB::kan$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP108082	this study
LL1269	$\Delta hfsC::kan$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP108085	this study
LL1270	$\Delta hfsD::kan$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP108087	this study
LL1349	$\Delta hsfAB::kan$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP108105	this study
LL1350	$\Delta hfsCD::kan$	<i>Thermoanaerobacterium saccharolyticum</i>	SRP108109	this study
LL1244	<i>wt</i>	<i>Thermoanaerobacterium thermosaccharolyticum</i>	SRP096459	DSMZ 571
LL1346	$\Delta hfsB::kan$	<i>Thermoanaerobacterium thermosaccharolyticum</i>	SRP108099	this study
LL1301	<i>wt</i>	<i>Thermoanaerobacterium xylanolyticum</i>	SRP108095	DSMZ 7097
LL1347	$\Delta hfsB::kan$	<i>Thermoanaerobacterium xylanolyticum</i>	SRP108102	this study
LL1258	<i>wt</i>	<i>Thermoanaerobacter mathranii</i>	SRP108092	DSMZ 11426
LL1348	$\Delta hfsB::kan$	<i>Thermoanaerobacter mathranii</i>	SRP108103	this study
LL1469	$\Delta hfsA::kan \Delta ldh::ery$	LL1267 ( <i>T. saccharolyticum</i> )		this study
LL1470	$\Delta hfsB::kan \Delta ldh::ery$	LL1268 ( <i>T. saccharolyticum</i> )		this study
LL345	$\Delta hpt$	<i>Clostridium thermocellum</i>	SRP053786	[119]
LL1048	$\Delta hfs::P_{gapD-cat}$	<i>Clostridium thermocellum</i>	SRP108090	this study
LL1474	$\Delta hpt \Delta hfsB$	LL345 ( <i>Clostridium thermocellum</i> )		this study
LL1475	pDGO143(P <sub>eno</sub> - <i>hfs</i> )	LL1147 ( <i>Clostridium thermocellum</i> )		this study
LL1476	pDGO143(P <sub>2638</sub> - <i>hfs</i> )	LL1147 ( <i>Clostridium thermocellum</i> )		this study
LL1477	pDGO143(P <sub>hfs</sub> - <i>hfs</i> )	LL1147 ( <i>Clostridium thermocellum</i> )		this study
LL1478	$\Delta hpt \Delta hydG \Delta ech$ pDGO143 (empty)	LL1147 ( <i>Clostridium thermocellum</i> )		this study
LL1147	$\Delta hpt \Delta hydG \Delta ech$	LL345 ( <i>Clostridium thermocellum</i> )		[128]
pDGO145	Deletion of <i>hfsB</i> deletion in LL345		KY852359	[143]
pDGO143	<i>C. thermocellum</i> expression plasmid		KX259110	[144]
pAEHFS1	Plasmid pDGO143 expressing <i>T. saccharolyticum hfsABCD</i> genes under controls of the P <sub>eno</sub> promoter			this study



Continuous of Table 1.

pAEHFS2	Plasmid pDGO143 expressing <i>T. saccharolyticum</i> <i>hfsABCD</i> genes under controls of the P <sub>2638</sub> promoter		this study
pAEHFS3	Plasmid pDGO143 expressing <i>T. saccharolyticum</i> <i>hfsABCD</i> genes under controls of the P <sub>hfs</sub> promoter		this study

*Thermoanaerobacterium saccharolyticum*, *Thermoanaerobacterium xylanolyticum* LX-11 and *Thermoanaerobacterium thermosaccharolyticum* were grown anaerobically with an initial pH of 6.2 (*T. saccharolyticum* and *T. xylanolyticum*) and 6.7 (*T. thermosaccharolyticum*) and 7.0 (*Clostridium thermocellum*) at 55°C.

*Thermoanaerobacter mathranii* subsp. *mathranii* str. A3 was grown anaerobically at 60 °C with initial pH 7.4.

For transformations *T. saccharolyticum*, *T. xylanolyticum*, *T. thermosaccharolyticum* and *T. mathranii* were grown in modified DSMZ M122 rich media (pH: 6.7 for kanamycin, pH: 6.1 for erythromycin selections) contains per liter 10.0 g xylose, 3 g Na<sub>3</sub>C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>·2H<sub>2</sub>O, 2.6 g MgCl<sub>2</sub>·6H<sub>2</sub>O, 1.5 g KH<sub>2</sub>PO<sub>4</sub>, 0.13 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.00013 g FeSO<sub>4</sub>·7H<sub>2</sub>O, 4.5 g yeast extract, 0.5 mL 0.2% wt/vol Resazurin, 0.5 g L-cysteine-HCl, 10 g MOPS sodium salt, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and 1.2% agarose for solid media at 55 °C. as previously described [104]. For erythromycine selections, incubation was performed at 48 °C.

Transformation of *C. thermocellum* were carried out using the modified DSM122 medium designated as CTFUD, contains per liter 5.0 g cellobiose, 3 g Na<sub>3</sub>C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>·2H<sub>2</sub>O, 2.6 g MgCl<sub>2</sub>·6H<sub>2</sub>O, 1.5 g KH<sub>2</sub>PO<sub>4</sub>, 0.13 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.00013 g FeSO<sub>4</sub>·7H<sub>2</sub>O, 4.5 g yeast extract, 0.5 mL 0.2 % wt/vol Resazurin, 0.5 g L-cysteine-HCl, 10 g MOPS sodium salt, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and 0.8 % agarose for solid media (pH: 7.0) at 55 °C.

8-Azahypoxanthine (8AZH) selection was carried out in NY-CTFUD medium contains 5.0 g per liter cellobiose, 3 g Na<sub>3</sub>C<sub>6</sub>H<sub>5</sub>O<sub>7</sub>·2H<sub>2</sub>O, 2.6 g MgCl<sub>2</sub>·6H<sub>2</sub>O, 1.5 g KH<sub>2</sub>PO<sub>4</sub>, 0.13 g CaCl<sub>2</sub>·2H<sub>2</sub>O, 0.00013 g FeSO<sub>4</sub>·7H<sub>2</sub>O, 1.0 mL vitamin solution, 0.5 mL 0.2 % wt/vol Resazurin, 0.5 g L-cysteine-HCl, 10 g MOPS sodium salt, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and 0.8 % agarose for solid media (pH: 7.0) at 55 °C.

All inoculations and transformations (except the *E. coli* clonings) were performed in an anaerobic chamber (COY Laboratory Products, Inc., Grass Lake, MI, USA) at 85% N<sub>2</sub>, 10% CO<sub>2</sub>, and 5% H<sub>2</sub> atmosphere under < 5 ppm oxygen were maintained using a palladium catalyst.

## 2.4. Bioinformatic Analysis and Determination of the Microorganisms to be Studied

Whole genome sequence of *T. saccharolyticum* JW/SL-YS485 have been deposited in NCBI under the accession number CP003184.1. This sequence was retrieved from the NCBI and used for bioinformatic analysis and homologous recombination primer design. The complete nucleotide sequence GeneBank accession number of the *hfs* operon is GQ354412. The *hfsA*, *hfsB*, *hfsC* and *hfsD* amino acid sequences were also imported from NCBI with the accession numbers ACU11594.1, ACU11595.1, ACU11596.1, ACU11597.1 respectively.

In order to see whether disruption of the *hfsB* gene affect ethanol production similarly in other microorganisms, some candidate microorganisms were identified. For this purpose organisms including a possible *hfs* operon with *nfnAB* and *adhA* genes in their genomes were searched from the NCBI's nucleotide collection (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) using the nucleotide sequences of the *T. saccharolyticum* JW/SL-YS485 *hfs*, *nfnAB* and *adhA* genes. Among the listed organisms *Thermoanaerobacterium xylanolyticum* LX-11, *Thermoanaerobacterium thermosaccharolyticum* and *Thermoanaerobacter mathranii* subsp. *mathranii* str. A3 were selected as candidate strains for *hfsB* inactivation.

*C. thermocellum*'s and *T. mathranii*'s *hfs* operon structure are different from other organism, thus it was decided to use to evaluate the difference.

Protein domain analysis was performed using the NCBI's Conserved Domain Database (<https://www.ncbi.nlm.nih.gov/cdd>).

## 2.5. Primers

Primers used in this study were listed in Table 2. For isothermal gibson assembly, primers were designed to contain 15-20 bp overlaps to the adjacent sequence that target of the assembly.

Table 2. List of primers used in this study

<b><i>T. saccharolyticum</i> <math>\Delta</math><i>hfsA</i>::<i>kan</i> Primers</b>
<i>hfsA</i> Up stream flanking fragment

**Continuous of Table 2.**

AE2: 5'-CATGGCGTCGGGTCTTCTGTC-3'

AE3: 5'-TGGGGTAAACATTTGTCAAGAATCCGTATACTGTATAC-3'

hfsA kan fragment

AE1: 5'-ATTCTTGACAAATGTTTACCCCATTAG -3'

AE4: 5'-GCTTCTTTGAAATTAATGACTCATATCCGATACAAATTCCTCGTAG-3'

hfsA Down stream flanking fragment

AE5: 5'-ATGAGTGTCAATTAATTTCAAAGAAGC-3'

AE6: 5'-GTA ATG CCG CTG AAA TCA CCG -3'

hfsA External Confirmation

AE24: 5'-TATTCTTTTCCACTTAAAGCTGTAG-3'

AE25: 5'- TGC CAA TCT TCG ACA ATA TCT C -3'

***T. saccharolyticum*  $\Delta$ hfsB::kan Primers**hfsB Up stream flanking fragment

AE7: 5'-GGTCTTCCTGTCGTAGCAACAG-3'

AE8: 5'-GTAAACATTTGTCAAGAATCATTTTAGTTCACC-3'

hfsB kan fragment

AE1: 5'-ATTCTTGACAAATGTTTACCCCATTAG-3'

AE9: 5'-GCAATATCGATGTAATGACTCATATCCGATACAAATTCCTC-3'

hfsB Down stream flanking fragment

AE10: 5'- ATGAGTCATTACATCGATATTGCACATG-3'

AE11: 5'-CACAATCCTGACCTTTATTCTAAGATCG-3'

hfsB External confirmation

AE26: 5'-ATGTTCTTAGGAAATTATCGAGAG-3'

AE27: 5'-CAT TGCGACTTTAGTTCTCTGC-3'

***T. saccharolyticum*  $\Delta$ hfsC::kan Primers**hfsC Up stream flanking fragment

AE12: 5'- GAGAATTTAGAAGATGTTGACTCTG-3'

AE13: 5'-GGTAAACATTTGTCAAGAATCATTCATCATCACCTG-3'

hfsC kan fragment

AE13: 5'- ATTCTTGACAAATGTTTACCCCATTAG-3'

AE14: 5'- ACAAGGCAATCTATCCTTTATCCGATACAAATTCCTCGTAG-3'

hfsC Down stream flanking fragment

AE15: 5'- AAAGGATAGATTGCCTTGTTTAAG-3'

<b>Continuous of Table 2.</b>
AE16: 5'- GCATCCAACACAGCAACCATCTC-3'
<u>hfsC External confirmation</u>
AE28: 5'-CTGTGAAGGCGCTATAGATGC-3'
AE29: 5'- TCC ACT TCC ACA TCT TGC TG -3'
<b><i>T. saccharolyticum</i> <math>\Delta</math>hfsD::<i>kan</i> Primers</b>
<u>hfsD Up Stream flanking fragment</u>
AE17: 5'- TGGATATGGGTTCTGAGCTAAGAG-3'
AE18: 5'- CTAATGGGGTAAACATTTGTCAAGAATCTGAACATCTGTG-3'
<u>hfsD kan fragment</u>
AE1: 5'- ATTCTTGACAAATGTTTACCCCATAG-3'
AE19: 5'- GAT GGA CCC ATT ATT TTT AAA ACT CAT CCG ATA CAA ATT CC -3'
<u>hfsD Down stream flanking fragment</u>
AE20: 5'-GAGTTTTAAAATAATGGGTCCATC-3'
AE21: 5'-ATATTTGAGGAGCACCGTCGTC-3'
<u>hfsD External confirmation</u>
AE30: 5'- CTACTTTGACAACGCGAATAGTG-3'
AE31: 5'-GCAGCACCTCACCTCTTACAC-3'
<b><i>kanR</i> Gene Internal Primers</b>
AE22: 5'-GCACTTTGAACGGCATGA TGG -3'
AE23: 5'- CCATCATGCCGTTCAAAGTGC-3'
<b><i>T. xylanolyticum</i> <math>\Delta</math>hfsB::<i>kan</i> Primers</b>
<u>hfsB Up stream flanking fragment</u>
AE51: 5'-TACGGACGAAACGGAAGGGAATTTGC-3'
AE52: 5'-GGGGTAAACATTTGTCAAGAATTTTAGTTCACCTAAAACATATTC-3'
<u>hfsB kan fragment</u>
AE1: 5'-ATTCTTGACAAATGTTTACCCCATAG-3'
AE53: 5'-GCAATATCGATGTAATGACATCCGATACAAATTCCTCGTAG-3'
<u>hfsB Down stream flanking fragment</u>
AE5: 5'-GATGTCATTACATCGATATTGC-3'
AE55: 5'-CCACTAAGAAATTTACGTCAGTAGCG-3'
<u>hfsB External onfirmation</u>
AE56: 5'-GCAACAGACATTGGCGATTTAGG-3'

<b>Continuous of Table 2.</b>
AE57: 5'-AGCTCTCCACAATCCTGACC-3'
<b><i>T. thermosaccharolyticum</i> <math>\Delta hfsB::kan</math> Primers</b>
<u><i>hfsB</i> Up stream flanking fragment</u>
AE61: 5'-GAGATGAGCATAATCTTTGAAAACTGG-3'
AE62: 5'-GGTAAACATTTGTCAAGAATTGTAAATTACCTAGCACATATTCTC-3'
<u><i>hfsB</i> kan fragment</u>
AE1: 5'-ATTCTTGACAAATGTTTACCCATTAG-3'
AE58: 5'-CAATATCAATGTAATGGCTCATATCCGATACAAATTCCTCG-3'
<u><i>hfsB</i> Down stream flanking fragment</u>
AE63: 5'-ATGAGCCATTACATTGATATTGCAC-3'
AE64: 5'-GATCCGCAGGAAAATCAGGATTTGATG-3'
<u><i>hfsB</i> External confirmation</u>
AE73: 5'-CGTGATAAAATGAGTATATACG-3'
AE74: 5'-CTCTAATAAGTTAATAAGTTCTTCAAC-3'
<b><i>Thermoanaerobacter mathranii</i> subsp. <i>mathranii</i> str. A3 <math>\Delta hfsB::kan</math> Primers</b>
<u><i>hfsB</i> Up stream flanking fragment</u>
AE65: 5'-CACCTTACAGCATTCTATTGG-3'
AE66: 5'-GGTAAACATTTGTCAAGAATAATTATTCCTCCTCCCATTAC-3'
<u><i>hfsB</i> kan fragment</u>
AE1: 5'-ATTCTTGACAAATGTTTACCCATTAG-3'
AE59; 5'-CTTCAGCAAAGACCTTCATATCCGATACAAATTCCTC-3'
<u><i>hfsB</i> Down stream flanking fragment</u>
AE67: 5'-ATGAAGGTCTTTGCTGAAGTGTTAC-3'
AE68: 5'-GTTGAGCGATTGACAAGCAGG-3'
<u><i>hfsB</i> External confirmation</u>
AE75: 5'-GAGTGCCTCTATCCTATGTATGC-3'
AE76: 5'-GCCAGCTCTTCTACTACCTTGAG-3'
<b><i>T. saccharolyticum</i> <math>\Delta hfsAB::kan</math> primers</b>
<u><i>hfsAB</i> Up stream flanking fragment</u>
AE2: 5'-CATGGCGTCGGGTCTTCCTGTC-3'
AE3: 5'-TGGGGTAAACATTTGTCAAGAAT CCGTATACTGTATAC-3'
<u><i>hfsAB</i> kan fragment:</u>

**Continuous of Table 2.**

AE1: 5'-ATTCTTGACAAATGTTTACCCCATTAG-3'

AE9: 5'- GCAATATCGATGTAATGACTCATATCCGATACAAATTCCTC-3'

hfsAB Down stream flanking fragment:

AE10: 5'-ATG AGT CAT TAC ATC GAT ATT GCA CAT G -3'

AE11: 5'-CACAATCCTGACCTTTATTCTAAGATCG-3'

hfsAB external confirmation

AE24: 5'-TATTCTTTTCCACTTAAAGCTGTAG-3'

AE27: 5'-CAT TGC GAC TTT AGT TCT CTG C-3'

***T. saccharolyticum*  $\Delta$ hfsCD::kan primers**hfsCD Up stream flanking fragment

AE12: 5'- GAGAATTTAGAAGATGTTGACTCTG -3'

AE13: 5'-GGTAAACATTTGTCAAGAATTCATTCATCATCACCTG-3'

hfsCD kan fragment:

AE1: 5'- ATTCTTGACAAATGTTTACCCCATTAG-3'

AE18: 5'- CTAATGGGGTAAACATTTGTCAAGAATCTGAACATCTGTG-3'

hfsCD Down stream flanking fragment:

AE20: 5'-GAGTTTTTAAAAATAATGGGTCCATC-3'

AE21: 5'-ATATTTGAGGAGCACCGTCGTC-3'

hfsCD External confirmation

AE28: 5'-CTGTGAAGGCGCTATAGATGC-3'

AE31: 5'-GCAGCACCTCACCTCTTACAC-3'

***T. saccharolyticum*  $\Delta$ hfsA::kan $\Delta$ ldh::erm Primers**ldh Up stream flanking fragment

AE32: 5'-ATCGAGGTATCCAAGCGATTCAATAG-3'

AE33: 5'-GTTTCGCTGGGTT TAT CGA CCG CTT TAG CTA TCA TGC CTT AGC-3'

erm fragment

AE34: 5'- GGTCGATAAACCCAGCGAACC -3'

AE35: 5'-CTCGATATAATCTCTTGCGGGCTGTTATGCTCCATGCTGCAAACCTCTG-3'

ldh Down stream fragment

AE36: 5'- AGC CCG CAA GAG ATT ATA TCG AG -3'

AE37: 5'- GTG ATG TAA GGA ACT ATA AGC GC -3'

External and internal confirmation

**Continuous of Table 2.**

AE38: 5'-GCT CAT GAA CCC AAA GTT GCA AAG C-3'

AE39: 5'-CTC CTG CAT TGC CTA CAA AGT ACA -3'

AE40: 5'- AGA CAG TCA TCT ATT CAA CTT ATC G -3'

***C. thermocellum*  $\Delta$ hfsB Primers**5' Flanking fragment

AE317: 5'-AGGCGTATCACGAGGCgatCCAGCAACTATTCAAAGGCC

AE318: 5'-TAGACTCGCGCTCCTTATC-3'

3' Flanking fragment

AE319: 5'-GATAAGGAGCGCGAGTCTAATGAATGATTTATGCGTTG-3'

AE320: 5'-CATGcCTATTTCCACgatCATTCTTGCTATTATCGAAG-3'

Internal region fragment

AE323: 5'-CCTGGCCCAGTAGTTcagGCTTGGGTTTAAAGGTGCGG-3'

AE324: 5'-TTCACTACTATTAGcagAGACGGGCTTAAGTTC-3'

Vector backbone I

AE321: 5'-atcGTGGGAATAGgCATG-3'

AE322: 5'-ctgAACTACTGGGCCAGG-3'

Vector backbone II

AE315: 5'-ctgCTAATAGTAGTGAAAAAATC-3'

AE316: 5'-atcGCCTCGTGATACGCCT-3'

5' flank-3' flank confirmation for pDGO145-hfsB in *E. coli*

AE115: 5'-CACCTGACGTCTAAGAAA-3'

AE118: 5'-TCTTTTCTCTCTTTTCGG-3'

Internal region confirmation for pDGO145-hfsB in *E. coli*

AE119: 5'-TTTAAACCCGCTGATCCT-3'

AE120: 5'-GTTGTCTAACTCCTTCCT-3'

Marker Integritycat-hpt

AE198: 5'-GCTATCTTTACAGGTACATCATTCTGTTTGTG-3'

AE199: 5'-TTTCATCAAAGTCCAATCCATAACCC-3'

cbp-tdk

AE200: 5'-ACTTCATGGCACTTTCTACACCTTGC-3'

AE201: 5'-TCGGAGTAAGGTGGATATTGATTTGC-3'

**Continuous of Table 2.**Chromosomal Confirmation (External)

AE222: 5'-GTCTTGCCCATTTGCTTAAG-3'

AE223: 5'-CGC TTA TGA ATG AAA TCA GGC-3'

Chromosomal Confirmation (Internal)

AE224: 5'- CTG CTA ACA GGA CTT CCC -3'

AE225: 5'-GTGCTGCCATACGGTAAAC-3'

***T. saccharolyticum adhA and adhE expression****T. saccharolyticum recA* qPCR primers

gblock:

GTTGAAGCCTTAGTGCGAAGTGGTGCTGTGGATGTGATCGTTATTGACTCTGTAGCTGCTC  
TCGTACCGAAAGCAGAGATAGATGGTGATATGGGCGATGCACATGTTGGACTTCAAG

recA qPCR F: 5'-GAAGCCTTAGTGCGAAGTGG-3'

recA qPCR R: 5'-GAAGTCCAACATGTGCATCG-3'

*T. saccharolyticum adhA* qPCR primersAATAGCTCATGGTTTAGGGCTTGGTGCAATATTGCCAGCAGTTATAAAAAGCTATTTATCCA  
GCTACAGCAGAAGTATTGGCTGATGTATATAGTCCTATAGTTCCTGGTTTAAAAGGACTGC  
CTGTTGAGGCGGAGTATGTAGC

adhA qPCR F: 5'-AGCTCATGGTTTAGGGCTTG-3'

*T. saccharolyticum adhA* qPCR R: 5'-ACATACTCCGCCTCAACAGG-3'*T. saccharolyticum adhE* qPCR primersTCAGCTTCATCCAAAGGCAATAAAGTGCAGCATCGCAGCAGCCAAAGTGATGTATGAAG  
CTGCACTAAAGGCAGGCGCACCTGAAGGATGCATAGGATGGATAGAAACGCCATCAATTG  
AGGCCA

adhE qPCR F: 5'-GCTTCATCCAAAGGCAATA-3'

adhE qPCR R: 5'-CCTCAATTGATGGCGTTTCT-3'

***C. thermocellum hfsABCD expression***pDGO143(P<sub>hfs</sub>-hfs)AE81; 5'- AAAGAGTAGTTCAACCCAAATTATTTATTAAAAGAGTAGTTCAACC-3'

AE50;

5'-CTCTAGAGGATCCCCGTTAATAAGGATATATATTTCAACAAAAATTTAG-3'



<p><b>Continuous of Table 2.</b></p> <p>pDGO143(P<sub>eno</sub>-<i>hfs</i>)</p> <p>AE44; 5'-AAAGAGTAGTTCAACCCGGAAATATTTAAAATGGAAATGTTGAAAAAATG-3'</p> <p>AE45; 5'-TAATAACCATTTCTCCATTCTCCCTTCATATAG-3'</p> <p>AE46; 5'-GAATGGAGAATGGTTATTACTGTTTGTGTAG-3'</p> <p>AE50; 5'-CTCTAGAGGATCCCCGTTAATAAGGATATATATTTCAACAAAAATTTAG-3'</p>
<p>pDGO143(P<sub>2638</sub>-<i>hfs</i>)</p> <p>AE47; 5'-AAAGAGTAGTTCAACCCGATAAACAAGGACGGTTC-3'</p> <p>AE48; 5'-TAATAACCATAACAAATTCCTCCTTACTTTTG-3'</p> <p>AE49; 5'-GGAATTTGTTATGGTTATTACTGTTTGTGTAG-3'</p> <p>AE50; 5'-CTCTAGAGGATCCCCGTTAATAAGGATATATATTTCAACAAAAATTTAG-3'</p>
<p><u>External confirmation of fragment integration</u></p> <p>AE98; 5'-GACGAAAAAGCCGATGAAGATGG-3'</p> <p>AE99; 5'-CCGTTTGTAAGTTAAAC-3'</p>

## 2.6. Molecular Studies

### 2.6.1. Construction of *hfsA*, *B*, *C*, *D* Knockout *Thermoanaerobacterium* and *Thermoanaerobacter* Strains

All knockout fragments were constructed via isothermal DNA assembly (Gibson Assembly) [145, 146]. Isothermal Gibson assembly is an *in vitro* recombination method that could assemble dsDNA in a single step.

Disruption of each gene were performed via allelic replacement by one step homologous recombination.

#### 2.6.1.1. Preparation of Gene Deletion Cassettes

Disruption of *hfsA*, *hfsB*, *hfsC* and *hfsD* genes in *T. Saccharolyticum* and deletion of *hfsB* gene in *T. thermosaccharolyticum*, *T. xylanolyticum* and *T. mathranii* was carried out with amplification of up and down stream flanking regions of each gene by PCR using primers given at Table 2. Kanamycin resistance gene from plasmid pMU433 [103] was also amplified by PCR and used as selection marker. For this purpose, each linear assembly

fragment was obtained by PCR using Phusion Hifi Master Mix in 50  $\mu$ L reaction volume containing 2  $\mu$ L cell culture or 1  $\mu$ L pure plasmid, 1X Phusion Hifi Master Mix, 10  $\mu$ M each primer with Bio-Rad S1000™ Thermal Cycler with Dual 48/48 Fast Reaction Modules (Bio-Rad, Hercules, CA, USA). For confirmation of amplifications, amplicons run on 1% agarose gel with 2-log Quick Load Purple DNA ladder (New England Biolabs) and visualized by Gel Doc™ XR+ Gel Documentation System (Bio-Rad). The cycling conditions for each fragment were given at Table 3.

**Table 3.** PCR reaction conditions for knockout fragment amplification

<i>T. saccharolyticum</i>		
<i>hfsA, hfsB, hfsC, hfsD KanR</i> Fragments		
98 °C	5 min	
98 °C	10 sec	
57 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	
<i>hfsA</i> Upstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
60.2 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	
<i>hfsA</i> Downstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
55.5 °C	30 sec	×34
72 °C	40 sec	
72 °C	10 min	
<i>hfsB</i> Upstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
60.5 °C	30 sec	×34
72 °C	40 sec	
72 °C	10 min	
<i>hfsB</i> Downstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
58.5 °C	30 sec	×34
72 °C	40 sec	
72 °C	10 min	

Continuous of Table 3.

<i>hfsC</i> Upstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
54.5 °C	30 sec	×34
72 °C	40 sec	
72 °C	10 min	
<i>hfsC</i> Downstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
54.5 °C	30 sec	×34
72 °C	40 sec	
72 °C	10 min	
<i>hfsD</i> Upstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
59 °C	30 sec	×34
72 °C	40 sec	
72 °C	10 min	
<i>hfsD</i> Downstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
54 °C	30 sec	×34
72 °C	40 sec	
72 °C	10 min	
<i>T. thermosaccharolyticum</i>		
<i>hfsB</i> Upstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
57.8 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	
<i>hfsB</i> Downstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
57 °C	30 sec	×34
72 °C	30 sec	
72 °C	10 min	
<i>KanR</i> Fragments		
98 °C	30 sec	
98 °C	10 sec	
57 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	

Continuous of Table 3.		
<i>T. xylanolyticum</i>		
<i>hfsB</i> Upstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
63.5 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	
<i>hfsB</i> Downstream Flanking Fragment		
98 °C	5 min	
98 °C	10 sec	
52.4 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	
<i>KanR</i> Fragment		
98 °C	30 sec	
98 °C	10 sec	
57 °C	30 sec	×34
72 °C	1.30 min	
72 °C	10 min	
<i>T. mathranii</i>		
<i>hfsB</i> Upstream Flanking Fragment		
98 °C	30 sec	
98 °C	10 sec	
54.1 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	
<i>hfsB</i> Downstream Flanking Fragment		
98 °C	30 sec	
98 °C	10 sec	
59.7 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	
<i>KanR</i> Fragment		
98 °C	30 sec	
98 °C	10 sec	
57 °C	30 sec	×34
72 °C	1 min	
72 °C	10 min	

After amplification, fragments purified by QIAquick PCR Purification Kit following the manufacturer's instruction and eluted with dH<sub>2</sub>O. After purification DNA concentrations

were measured by Nanodrop 1000 and fragments assembled using 1:1:1 ratio in 20  $\mu$ L NEB Hifi DNA Assembly Master Mix, at 55 °C for 20 min.

Each joined product was diluted to 1 : 30 with dH<sub>2</sub>O and 1  $\mu$ L was used as template for PCR using Phusion Hifi Master Mix in 50  $\mu$ l reaction volume, with the following conditions: initial denaturation steps at 98°C for 30 sec, denaturation at 98 °C for 10 sec, followed by annealing at 54-58.6 °C (depending on the Primers T<sub>m</sub>) for 30 sec and primer extension at 72 °C for 2.10-2.30 min (depending on the fragment size), followed by a step at 72 °C for 10 min, for 34 cycles. After confirmation of fragment sizes on 1% agarose gel, resulted linear DNA fragments were used for gene deletions.

### **2.6.1.2. Transformation of the Deletion Fragments**

Transformation of *Thermoanaerobacterium* and *Thermoanaerobacter* strains was performed in an anaerobic chamber with natural competency method [23] with the following steps;

1. Transformations started by inoculation of 10 mL CTFUD medium (10 g/L xylose pH. 6.7) with a 2  $\mu$ L frozen bacteria stock and mixed.
2. Then 1 mL of this mixture was transferred to the tubes containing 2  $\mu$ g of Gibson assembled linear DNA fragments.
3. 1 mL DNA-free tube was also prepared for negative control in order to check the contamination.
4. The control tube and the DNA containing tubes are then incubated at 55°C for overnight to an optical density of 0.5-0.6.
5. After the OD reached to an 0.5-0.6, kanamycin or erythromycin containing 30 mL of media poured equally into the quad divided petri dishes.
6. 1 mL of transformation mix, mixed with the first quad and 50  $\mu$ L removed from the first quad and mixed into the second quad.
7. Same procedure repeated for the 3<sup>rd</sup> and 4<sup>th</sup> quads and all petri dishes were allowed to solidify at room temperature for 30 minutes or until completely solid and incubated at 55°C in a moisture retaining container until colony formation (~48 hours or until colonies appear).

*T. saccharolyticum* colonies was selected on 20 µg/ mL erythromycin or on 200 µg/ mL kanamycin while *T. xylanolyticum* and *T. thermoscharolyticum* were selected on 500 µg/mL and *T. mathranii* was selected on 1000 µg/mL kanamycin included agar media.

After colony formation, gene disruptions confirmed by colony PCR by one external-external and internal-external primer pairs (Table 2.), using 12.5 µL OneTaq Quick-Load 2X Master Mix (New England Biolabs), with the following conditions: initial denaturation steps at 94°C for 5 min, denaturation at 94 °C for 30 sec, followed by annealing at 45-60 °C for 1 min and primer extension at 68 °C for 1 min 30 sec, followed by a step at 68 °C for 5 min, for 34 cycles. In some condition 5% DMSO also added into the reactions.

### 2.6.2. Deletion of the *hfsB* gene in *Clostridium thermocellum* DSM 1313

In this study, a genetically tractable  $\Delta hpt$  strain of *C. thermocellum* DSM1313 (referred as LL345), that allows to use of two-stage selection, counter-selection method with 8-azahypoxanthin was used as the parent strain for *hfsB* deletion [119]. Gene loci of *hpt* and *hfsB* are Clo1313\_2927 and Clo1313\_1795 respectively.

#### 2.6.2.1. Construction of the Gene Deletion Plasmid

Plasmid pDGO145 (GeneBank accession number: KY852359) was used for deletion fragment construction (Figure 9). The plasmid pDGO145 was designed to contain three ~1000 bp flanking regions that are homologous to the up stream (5' flank), down stream (3' flank) regions and an internal flanking region that are homologous to the *hfsB* gene on the *C. thermocellum* chromosome. Vector backbone also amplified by PCR using the pDGO145 as template. The fragments were amplified by PCR with 1x Phusion Hifi Master Mix. Conditions for fragment amplification were given at Table 4.

Table 4. PCR conditions for *C. tehrmocellum hfsB* deletion fragment amplification

5' Flanking		
98 °C	30 sec	
98 °C	10 sec	
56.6 C	30 sec	×34
72 °C	40 sec	

<b>Continuous of Table 4.</b>		
72 °C		10 min
<b>3' Flanking</b>		
98 °C		30 sec
98 °C		10 sec
63.2 °C		30 sec ×34
72 °C		40 sec
72 °C		10 min
<b>Internal</b>		
98 °C		30 sec
98 °C		10 sec
63.2 °C		30 sec ×34
72 °C		40 sec
72 °C		10 min
<b>Vector backbone I</b>		
98 °C		30 sec
98 °C		10 sec
55 °C		30 sec ×34
72 °C		1.2 min
72 °C		10 min
<b>Vector backbone II</b>		
98 °C		30 sec
98 °C		10 sec
56 °C		30 sec ×34
72 °C		2.30 min
72 °C		10 min

Plasmid was constructed via Gibson assembly using 3:3:3:2:1 (5' flank: 3' flank: internal: vector backbone I:vector backbone II) ratio at 50 °C for 1 h and transformed into NEB 5-alpha competent *E. coli* cells. For his purpose 50 µL component *E. coli* cells that leaved on ice for thawing for 10 minute, than mixed with 20-30 µg plasmid DNA and placed on ice for 10 minute again. After 30 second heat shock at 42 °C, the mixture was placed on ice for 5 minute and 950 µL room temperature SOC medium (NEB, B9020S) added into the plasmid containing competent cells. Than tubes were incubated at 37 °C for 1 hour and plated on to 100 µg/L carbenicillin containing LB agar plates. Plasmids harbouring the correct fragment sizes were selected by colony PCR. Reaction was carried out in 25 µL final volume containing 0,5 µM each primer, 1X OneTaq Quick-Load® 2X Master Mix with Standard Buffer. Verified plasmids were transformed into the T7 Express Competent *E. coli* cells.

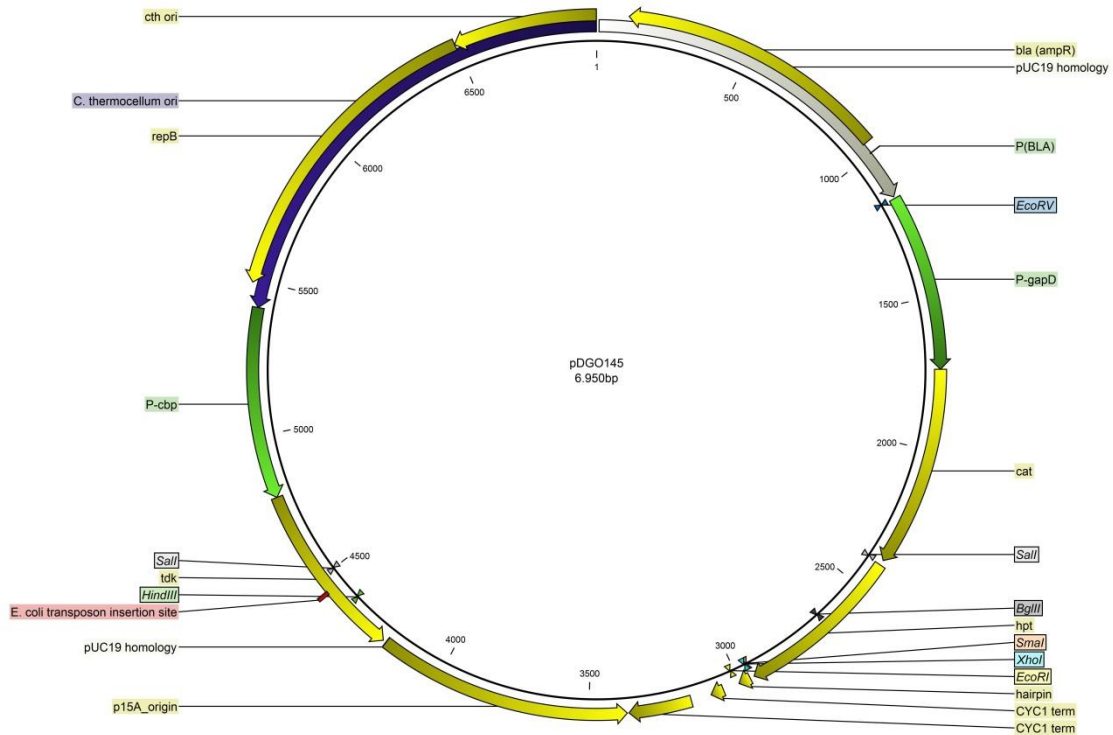


Figure 9. Map of the backbone vector pDGO145 was used for *hfsB* deletion in *C.thermocellum*

The plasmid containing the 5', 3' and internal regions of *hfsB* was isolated by the QIAprep Spin Miniprep Kit in accordance with the manufacturers procedure and transformed into LL345 via electroporation.

#### 2.6.2.2. Transformation of *hfsB*-pDGO145 into LL345 via Electroporation

Transformation of *hfsB*-pDGO145 into LL345 was performed by electroporation using the method described by Olson et. al. [9].

1. 50 mL CTFUD media was inoculated from an overnight culture of LL345 and incubated at 55 °C until OD<sub>600</sub>: 0.6–1.0 and tubes placed on ice for 20 min.
2. After 20 min on ice, cells were harvested by centrifugation at 6500 × g aerobically for 12 min at RT and supernatant was removed without disturbing the pellet.



3. Pellet was washed with 50 mL autoclaved reverse-osmosis-purified water (18 MΩ cmc) and tubes were centrifuged at  $6500 \times g$  for 12 min and supernatant was decanted.
4. Pellet was washed with the same conditions at step 3 and after centrifugation tubes were brought into the anaerobic chamber and cells resuspended with 100  $\mu$ L reverse-osmosis-purified water with pipetting carefully.
5. 20  $\mu$ L cell suspension and 150-200  $\mu$ g plasmid was added to a 1 mm electroporation cuvette and mixed.
6. 1500 V amplitude electrical pulse was applied for 1.5 ms (Bio-Rad Gene Pulser Xcell Microbial System part number 165-2662).
7. Then cells were resuspended in 3 mL CTFUD and incubated at 51 °C in a dry bath incubator for 16 h.
8. After incubation 20 mL CTFUD agar was melted and cooled down to 55 °C.
9. 6  $\mu$ g/mL thiamphenicol was added and mixed.
10. 1 mL and 50  $\mu$ L recovered cells were added into the melted agar by pipetting and poured into the petri dishes.
11. Poured petries was allowed to solidify at RT for 30 min and incubated in a second container at 55 °C until until colony formation observed.

### **2.6.2.3. Colony Selection and Marker Removal**

1. Following the observation of colony formation, colonies were selected and grown at 55°C in liquid CTFUD medium with 6  $\mu$ g/mL thiamphenicol and presence of plasmid was verified by PCR in 25  $\mu$ L final reaction volume with OneTaq Quick-Load 2X Master Mix
2. 1 mL plasmid containing cell culture was added into the 30 mL melted CTFUD agar supplemented with 6  $\mu$ g/mL thiamphenicol and 10  $\mu$ g/mL FUDR and poured into the quad divided petri dishes with  $4 \times 100$  serial dilutions.
3. Solidified petries were incubated at 55 °C until colony formation (3-4 days).
4. After colony formation, colonies picked into CTFUD-NY medium, incubated at 55 °C without selection and first and second recombination events were confirmed by PCR as described above.

5. Verified colony cultures were mixed with 30 mL molten CTFUD-NY including 500 µg/mL 8AZH.
6. Single colony was picked into liquid CTFUD medium and complete marker removal was confirmed with PCR.
7. For final confirmation and selection of pure colonies, the selected colony culture was grown on CTFUD agar and the single colony from this plate was used as a mutant strain.

### 2.6.3. Deletion of *hfsAB* and *hfsCD* in *Thermoanaerobacterium saccharolyticum* Strains

*hfsAB* and *hfsCD* knockout fragments were constructed via isothermal DNA assembly (Gibson Assembly).

As mentioned earlier, for this purpose each fragment to be used in the assembly were amplified with PCR with Phusion Hifi Master Mix in 50 µL reaction volume containing 2 µL cell culture, 1× Phusion Hifi Master Mix, 0,5 µM each primer. Reaction conditions for amplification of each fragment were given in Table 5.

Table 5. PCR reaction conditions for *hfsAB* and *CD* assembly fragment amplification

<b><i>hfsAB</i> Up Fragment (AE2-AE3)</b>	
98 °C	30 sec
98 °C	10 sec
63.3 °C	30 sec
72 °C	40 sec
72 °C	10 min
<b><i>hfsAB, kan</i> Fragment (AE1-AE9)</b>	
98 °C	30 sec
98 °C	10 sec
57 °C	30 sec
72 °C	1 min
72 °C	10 min
<b><i>hfsAB</i> Down Fragment (AE10-AE11)</b>	
98 °C	30 sec
98 °C	10 sec
58 °C	30 sec
72 °C	40 sec
72 °C	10 min
<b><i>hfsCD</i> Up Fragment (AE12-AE13)</b>	
98 °C	30 sec
98 °C	10 sec

54 °C	30 sec
72 °C	40 sec
72 °C	10 min
<b><i>hfsCD, kan</i> Fragment (AE1-AE18)</b>	
98 °C	30 sec
98 °C	10 sec
57 °C	30 sec
72 °C	1 min
72 °C	10 min
<b><i>hfsCD</i> Down Fragment (AE20-AE21)</b>	
98 °C	30 sec
98 °C	10 sec
53.7 °C	30 sec
72 °C	40 sec
72 °C	10 min

Size verified fragments were assembled using 1:1:1 ratio in 20 µL NEB Hifi DNA Assembly Master Mix, at 55 °C for 20 min. For transformation *hfsAB* and *hfsCD* deletion fragments were generated by PCR using the primers AE2-AE11 and AE12-AE21, respectively and used for gene deletion with natural competency method as previously mentioned.

Deletion of *hfsAB* (external AE24-AE27, internal AE24-AE27) and *hfsCD* (external AE28-AE31, internal AE28-AE23) were confirmed by colony PCR.

#### **2.6.4. Disruption of Lactate Dehydrogenase (*ldh*) Gene in *T. saccharolyticum* $\Delta hfsA$ (LL1267) and $\Delta hfsB$ (LL1268) Mutant Strains**

The lactate dehydrogenase (*ldh*) gene was inactivated in LL1267 and LL1268 using the same methods as described for *hfs* genes disruption, except *ery* resistant gene was used as the selection marker. First *ldh* up-down fragments were amplified by PCR with Phusion Hifi Master Mix in 50 µL final reaction volume containing 2 µL LL1025 cell culture, 1X Phusion Hifi Master Mix and 0.5 µM each primer with the conditions given at Table 6. Erythromycin resistance gene amplified from plasmid pMU620 [27] Linear gene deletion fragment was constructed by Gibson assembly as described above and transformed by natural competency.

Colony verification was carried out by colony PCR.

Table 6. PCR reaction conditions for *ldh* assembly fragments

<b><i>ldh</i> Up Fragment</b>	
98 °C	30 sec
98 °C	10 sec
59.4 °C	30 sec
72 °C	40 sec
72 °C	10 min
<b><i>ldh</i> Down Fragment</b>	
98 °C	30 sec
98 °C	10 sec
56.5 °C	30 sec
72 °C	40 sec
72 °C	10 min
<b><i>EryR</i> Fragment</b>	
98 °C	30 sec
98 °C	10 sec
61.2 °C	30 sec
72 °C	1 min
72 °C	10 min

### 2.6.5. Whole Genome Sequence Analysis

To confirm complete gene disruption and to investigate possible mutations that may be associated with genetic inactivation, genome of all of the mutant *Thermoanaerobacter* and *Thermoanaerobacterium* strains LL1267, LL1268, LL1269, LL1270, LL1346, LL1347, LL1348, LL1350 were resequenced as described previously [147]. Parent strains LL1025, LL1244, LL1301 and LL1258 were also resequenced for comparison.

#### 2.6.5.1. Bacterial Genomic DNA Isolation With CTAB

Genomic DNAs were isolated using the CTAB method [148] Bacteria were grown to OD<sub>600</sub>: 0.5-0.7 (mid log) in 250 mL CTFUD medium and cells were collected by centrifugation at maximum speed for 15 min. Then DNA isolation process was started.

1. Pelleted cells resuspended in TE buffer and OD<sub>600</sub> was adjusted to ~1.0 (1.5 mL).
2. Cell suspension was transferred to a clean centrifuge tube and 20 µL (100mg/mL) lysozyme was added and mixed.
3. Tubes were incubated at 37 °C for 30 min.
4. 40 µL 10 % SDS was added and mixed.

5. 8  $\mu$ L proteinase K (10mg/mL) was added and mixed and tubes were incubated at 56 °C for 16 h.
6. 5 M NaCl was added and mixed.
7. 100  $\mu$ L CTAB/NaCl that heated to 65 °C was added and incubated at 65 °C for 10 min.
8. 0.5 mL chloroform:isoamyl alcohol was added (24:1) and tubes were centrifuged at room temperature at maximum speed for 10 min.
9. Aqueous phase transferred into clean microcentrifuge tubes and phenol:chloroform:isoamyl alcohol (25:24:1) was added, mixed and centrifuged at maximum speed for 10 min at room temperature.
10. Aqueous phase was transferred into clean microcentrifuge tube and 0.5 mL chloroform:isoamyl alcohol (24:1) was added and tubes centrifuged again.
11. New aqueous phase was transferred and 0.6 vol isopropanol (-20 °C) was added and incubated at -20 °C for overnight.
12. After incubation tubes were centrifuged at maximum speed for 15 min at 4 °C.
13. Pellets were washed with cold 70 % ethanol and centrifuged at maximum speed for 5 min.
14. Supernatant discarded and pellet was left to dry at room temperature.
15. DNA was resuspended in ~170  $\mu$ l of DNase-free water.
16. 30  $\mu$ L RNase was added, mixed and incubated at 37 °C for 1 h.
17. Enzyme was inactivated at 70 °C for 15 min.
18. Tubes were cooled on ice and 1/10 volume of 3M Sodium Acetate and 2.5 volumes of 100% ethanol was added.
19. Samples were spinned down and were kept at -80°C for 30 min.
20. End of the incubation time sample tubes were centrifuged at 4°C for 20 min to pellet DNA.
21. Pellet was washed with 70 % cold ethanol and centrifuged again at 4 °C for 3-5 min.
22. Dried pellet was resuspended in 100  $\mu$ l of TE and stored at -80 °C until sequencing.

### **2.6.5.2. Whole Genome Sequence Analysis**

The sequence analysis was conducted in the Department of Energy Joint Genome Institute (JGI, Walnut Creek, CA, USA) with an Illumina MiSeq instrument as described previously [147]. Reads were subjected to the BLAST search [149].

### **2.6.6. Differential Expression Analysis of Mutant *T. saccharolyticum* Strains by RNA Sequencing**

For transcriptome analysis of each strain, RNA-seq and data normalization of sequenced libraries were carried out by DOE JGI (Department of Energy Joint Genome Institute). For this purpose, 20 ml RNA Protect Bacteria Reagent was mixed with 10 mL mid-log phase (OD<sub>600</sub> 0.1–0.2 for  $\Delta hfsA$  (LL1267),  $\Delta hfsC$  (LL1269) and  $\Delta hfsD$  (LL1270), 0.3–0.4 for  $\Delta hfsB$  (LL1268) and LL1049 and 0.6–0.8 for LL1189 and LL1025) bacterial cell culture (in MTC-6 defined medium) and was centrifuged. Total RNAs were extracted by RNeasy mini kit (Qiagen, CA, USA) following the manufacturer's instructions. Contaminated DNA was removed using RNase-Free DNase set (Qiagen, CA, USA).

The quality control of RNA samples to evaluate the quantity, performed by a Qubit 2.0 fluorometer using Life Technologies Quant-iT™ RNA Assay Kit. A three-point (50 ng/μL, 400 ng/μL, and 1000 ng/μL) standard curve along with a blank, was used for the assay. Agilent 2100 BioAnalyzer was used to analyze the quality. Purity of RNA was determined with NanoDrop Spectrophotometer. Data were normalized as described previously [150].

To determine genes related with increased ethanol production, the high ethanol-yielding strains (LL1049, LL1267, and LL1268) were compared with low ethanol yielding strains (wt, LL1187, LL1269, and LL1270) using a two-tailed t test. Data are represented as a volcano plot [151].

### **2.6.7. Reverse Transcription Quantitative PCR (RT-qPCR) Expression Analysis for Determination of Expression Levels**

Determination of gene expression levels was performed as described previously [147]. Strains were cultured in 20 ml MTC-6 defined medium, and harvested in mid-log-phase (OD<sub>600</sub> 0.6–0.8); 10 ml aliquots of the cell cultures were immediately treated with 20 ml

RNA Protect Bacteria Reagent. After centrifugation, pellets were stored at  $-80^{\circ}\text{C}$  until RNA purification. RNA was extracted by RNeasy mini kit and treated with RNase-Free DNase set to ensure there was no DNA contamination. 500 ng of RNA was used to synthesize cDNA using the iScript cDNA synthesis kit. Expressions levels were measured by CFX96 qPCR system (Bio-Rad, Hercules, CA, USA) with SsoFast™ EvaGreen® Supermix.  $55^{\circ}\text{C}$  used as annealing temperature to determine expression levels of *adhE* and *adhA*, genes in *T. saccharolyticum* strains. For each gene expression was normalized to *recA* RNA. The purity of the RNA samples confirmed via cDNA synthesis in the presence and absence of reverse transcriptase. After a qPCR was conducted using *recA* primers to ensure only background levels were detected in the samples lacking reverse transcriptase. Standard curves were generated using a synthetic DNA template (gBlock, IDT, Coralville, IA, USA) containing the amplicons. Primers used for qPCR are listed in Table 2.

#### **2.6.8. Quantitative Proteomics Analysis of Mutant Strains by MS/MS**

Quantitative proteomics analysis of mutant strains, were performed at Oak Ridge National Laboratory (Oak Ridge, Tennessee, USA).

Wild-type and *Δhfs* strains were processed for LC-MS/MS analysis as previously described [152]. Briefly, cell pellets from 50 mL mid-log phase cell culture were re-suspended in SDS lysis buffer (4% SDS, 100 mM Tris-HCl, pH 8.0), boiled, sonicated, and crude protein lysate quantified via BCA assay. Two milligrams of crude protein were adjusted to 25 mM dithiothreitol (DTT), boiled, then precipitated with trichloroacetic acid (TCA). The protein pellet was washed with cold acetone then air dried before resuspension in 8 M urea, 100 mM Tris-HCl, 5 mM DTT, pH 8.0. Cysteines were then alkylated with 15 mM iodoacetamide (IAA) before digestion with sequencing-grade porcine trypsin (1:100 [w/w] overnight at room temperature after dilution to 4 M urea followed by another 1:100 [w/w] addition for 4 hr after dilution to 2 M urea). Digests were then salted and acidified and tryptic peptides collected via filtration through a 10 kDa MWCO spin column (Vivaspin 2; GE Healthcare).

Five micrograms of tryptic peptides were loaded onto a biphasic (reversed-phase [RP] and strong-cation exchange [SCX]) MudPIT back column, washed, and placed in-line with an in-house pulled nanospray emitter packed with 15 cm of RP resin as previously described [153]. Peptides were separated by a 2-step mini-MudPIT LC-MS/MS analysis employing 50

mM and 500 mM ammonium acetate salt cuts followed by standard, two-hour RP gradients. Eluting peptides were measured and sequenced in real time by a hybrid LTQ-XL Orbitrap mass spectrometer operating in data-dependent acquisition (DDA) mode. Relevant DDA parameters include: 1 full scan (15k resolution) followed by 10 MS/MS scans; isolation width = 2.2 m/z; CID energy = 35%; dynamic exclusion residence time and window = 30 s and 20 ppm, respectively; monoisotopic precursor selection = on; charge state screening = reject unassigned charges. Resulting MS/MS spectra were searched against the *Thermoanaerobacterium saccharolyticum* proteome database, concatenated with common contaminants and reversed sequences to assess false-discovery rates, with the Myrimatch v. 2.1 search algorithm [154]. Peptide-spectral matches (PSM) were scored, filtered (peptide-level FDR < 1 %), and assigned to proteins via IDPicker v. 3.0 [155]. Identified proteins were quantified by MS1 intensity via IDPicker's label-free quantitation method and normalized across runs to evaluate proteins that had marked differences in abundance due to specific deletions in genes of the *hfs* operon.

### **2.6.9. Heterologous Expression of *hfsABCD*\* Gene in *Clostridium thermocellum* Under Three Different Promoter**

#### **2.6.9.1. Plasmid and Strain Constructions**

*C. thermocellum*  $\Delta$ hydG $\Delta$ ech strain was previously described [128]. The plasmid pDGO143 (Figure. 10) was used as the backbone for the construction of the expression plasmids.



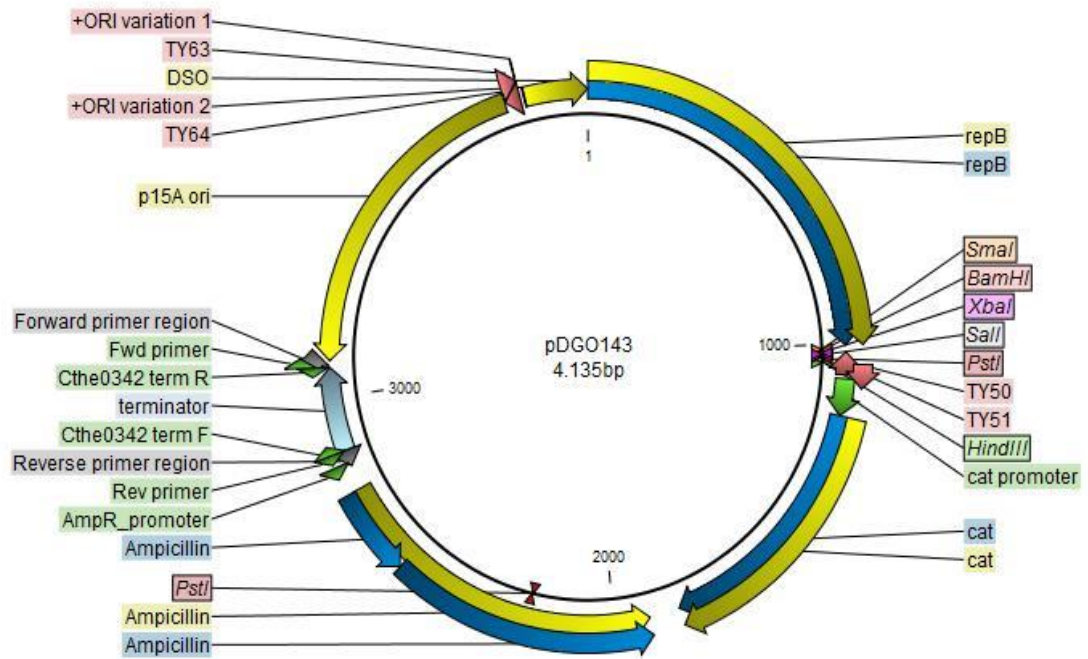


Figure 10. Map of the *hfsABCD* expression vector in *C. thermocellum*

The native *hfs* operon promoter sequence from *T. saccharolyticum* and the 2638 promoter and the *eno* promoter [10] sequences was used for the construction of the each expression vectors (Table 7).

Table 7. Description of promoter sequences used for *hfs* expression in *C. thermocellum*

Name	Length (bp)	Gene Function	Gene Locus
<i>eno</i>	178	glycolysis	Clo1313_2090
2638	209	peroxiredoxin	Clo1313_2638
<i>hfs</i>	357	hydrogenase	Tsac_1550-1553

The *hfs\** gene was amplified by PCR from the strain LL1049's genome with Phusion Hifi Master Mix in 50  $\mu$ l reaction volume. Cycling conditions of each fragments were given at Table 8.

Table 8. Fragment cycling conditions for *C. thermocellum hfsABCD* expression plasmid

<b><i>P<sub>eno</sub></i></b>	
98 °C	5 min
98 °C	10 sec
59 °C	30 sec
72 °C	15 sec
72 °C	10 min
<b><i>P<sub>eno-hfs</sub></i></b>	
98 °C	5 min
98 °C	10 sec
58 °C	30 sec
72 °C	2.5 min
72 °C	10 min
<b><i>P<sub>2638</sub></i></b>	
98 °C	5 min
98 °C	10 sec
59 °C	30 sec
72 °C	15 sec
72 °C	10 min
<b><i>P<sub>2638-hfs</sub></i></b>	
98 °C	5 min
98 °C	10 sec
58 °C	30 sec
72 °C	2.5 min
72 °C	10 min
<b><i>P<sub>hfs-hfs</sub></i></b>	
98 °C	5 min
98 °C	10 sec
57.4 °C	30 sec
72 °C	2.5 min
72 °C	10 min

Plasmids were constructed using standard molecular biology techniques and Gibson assembly [145, 146]. The expression vector was linearized by *Sma*I digested at RT for 1 h and linearized vector assembled with promoter and gene 3:2:1:1:1 ratio.

Expression plasmids were cloned in *E.coli* NEB5 cells and then transformed into T7 Express cells as described previously for  $\Delta hfsB$  plasmid construction. Colonies were verified by colony PCR as described above. Positive vectors were also verified by sequencing (Genewiz South Plainfield, NJ, USA).

#### **2.6.9.2. Transformation of the Expression Plasmids into *Clostridium thermocellum* $\Delta hydG\Delta ech$**

Plasmids from verified colonies were column purified and were transformed into *C. thermocellum*  $\Delta hydG\Delta ech$  by electroporation. Briefly, *C. thermocellum*  $\Delta hydG\Delta ech$  strain was grown overnight at 55 °C, inoculated into 50 mL CTFUD and the cell culture was centrifuged at 7000 rpm for 12 min at RT. The pellet then was washed twice with milliQ water and the resulted pellet was resuspended in 100  $\mu$ L ice chilled milliQ water. 20  $\mu$ L electrocompetent cell mixed with 300  $\mu$ g DNA and was added into the 1 mm electroporation cuvettes and 1500 V amplitude and 1.5 ms duration was applied.

Cells were recovered after 14-16 h incubation at 51 °C, in a dry bath incubator and mixed with the 6  $\mu$ g/ml Tm (thiamphenicol) supplemented melted CTFUD agar and poured into a petri dish and allowed to solidify. The petries incubated at 55 °C until colonies appear. Colonies were verified by colony PCR.

#### **2.6.9.3. Overexpressing the *hfsABCD* operon in $\Delta hydG\Delta ech$ *C. thermocellum* (LL1147)**

Strains were grown in MTC-5 minimal medium (5 g/L cellobiose, 9.3 g/L MOPS sodium salt, 2 g/L  $C_6H_5K_3O_7 \cdot H_2O$ , 1.3 g/L  $C_6H_8O_7 \cdot H_2O$ , 1.0 g/L  $Na_2SO_4$ , 1.0 g/L  $KH_2PO_4$ , 1.0 g/L  $NaHCO_3$ , 2.0 g/L  $CH_4N_2O$ , 1.0 g/L  $MgCl \cdot 6H_2O$ , 0.2 g/L  $CaCl_2 \cdot 2H_2O$ , 0.1 g/L  $FeCl_2 \cdot 6H_2O$ , 1.0 g/L L-Cysteine HCl monohydrate, 0.5  $\mu$ L trace minerals, 0.02 g/L pyridoxine HCl, 0.004 g/L 4-aminobenzoic acid, 0.002 g/L D-biotin, 0.002 g/L vitamin B12) supplemented by 6 g/L thiamphenicol at, 55°C for ~72 h.

## 2.7. Sample Collection and Analytical Studies

### 2.7.1. Bottle Fermentation and Analysis of the Fermentation Products

For fermentation analysis, *T. Saccharolyticum* (pH 6.2), *T. Xylanolyticum* (pH 6.2), *T. thermosaccharolyticum* (pH 7.0) and *T. mathranii* (pH 7.4) strains were grown in MTC-6 minimal medium containing 5 g/L or 20 g/L cellobiose, 9.3 g/L MOPS sodium salt, 2 g/L  $C_6H_5K_3O_7 \cdot H_2O$ , 1.3 g/L  $C_6H_8O_7 \cdot H_2O$ , 1.0 g/L  $Na_2SO_4$ , 1.0 g/L  $KH_2PO_4$ , 1.0 g/L  $NaHCO_3$ , 1.5 g/L  $NH_4Cl$ , 1.0 g/L  $MgCl \cdot 6H_2O$ , 0.2 g/L  $CaCl_2 \cdot 2H_2O$ , 0.1 g/L  $FeCl_2 \cdot 6H_2O$ , 1.0 g/L L-Cysteine HCl monohydrate, 0.5  $\mu$ L trace minerals, 0.02 g/L pyridoxine HCl, 0.004 g/L 4-aminobenzoic acid, 0.002 g/L D-biotin, 0.002 g/L vitamin B12, 0.004 g/L thiamine.

Fermentations were performed in 50 ml final volume in a 125 mL glass bottle sealed with a butyl rubber stopper [156]. The headspace was purged with nitrogen and the bottles were shaken at 200 rpm at 55°C for ~72 h.

#### 2.7.1.1. pH Measurements

Using a pH probe, the final pH of the bottles were measured by inserting a pH probe into the 15 mL tube containing the culture and recorded.

#### 2.7.1.2. Pressure

Headspace gas pressure in bottles was measured using a digital pressure gauge (Ashcroft, Stratford, CT, USA).

#### 2.7.1.3. Measurement of H<sub>2</sub> and CO<sub>2</sub> Percentage

Headspace H<sub>2</sub> and CO<sub>2</sub> percentage were measured using a gas chromatograph (Model; SRI Instruments, Torrance, CA, USA) with a HayeSep D packed column using a UV thermal conductivity detector with nitrogen carrier gas. Nitrogen flow rate was adjusted to approximately 8.2 mL/minute at 150°C. 50% H<sub>2</sub> and 25% CO<sub>2</sub> was used as standards.

#### 2.7.1.4. HPLC and TOCN (total organic carbon) analysis

Concentrations of cellobiose, glucose, lactate, acetate, ethanol, malate and formate were measured by high pressure liquid chromatography (HPLC) as previously described [6]. 700  $\mu$ L supernatant was run for analysis after acidified with 10% sulfuric acid ( $\text{H}_2\text{SO}_4$ ). Fermentation products in the liquid phase (cellobiose, xylose, ethanol, lactate, acetate, and formate) were measured using a Waters (Milford, MA, USA) HPLC with a HPX-87H column. Samples were run with a 1 $\times$ , 2 $\times$ , and 8 $\times$  diluted standards.

Pellet carbon and nitrogen, were used as an measurement of cell mass, and were measured with a Shimadzu TOC-V CPH elemental analyzer with TNM-1 and ASI-V modules (Shimadzu Corp., Columbia, MD, USA). After 3 washing step, the pellets were resuspended in 950  $\mu$ L milliQ water and centrifuged at maximum speed for 5 minutes at room temperature. 1 mL carbon/nitrogen pellet samples were added to TOCN tubes containing 19.5 mL MilliQwater. As carrier solvent, 8.72 mM HCl in distilled water was used. The resulting peak area was translated to TOCN values using the standard values in Microsoft Excel.

As standard solution, 1 mL glycine (5.00125 g/L) per tube and 8.333 mL of 6 M HCl per liter of distilled water was used in 1, 10, and 50 times dilution as well, along with several wash/ blank tubes containing MilliQ water.

#### 2.7.2. Fermentation Conditions for *C. thermocellum* strains

For batch fermentation, first mutant was grown for overnight at 55°C, in MTC-5 minimal medium containing 5 g/L cellobiose, 9.3 g/L MOPS sodium salt, 2 g/L  $\text{C}_6\text{H}_5\text{K}_3\text{O}_7 \cdot \text{H}_2\text{O}$ , 1.3 g/L  $\text{C}_6\text{H}_8\text{O}_7 \cdot \text{H}_2\text{O}$ , 1.0 g/L  $\text{Na}_2\text{SO}_4$ , 1.0 g/L  $\text{KH}_2\text{PO}_4$ , 1.0 g/L  $\text{NaHCO}_3$ , 2.0 g/L  $\text{CH}_4\text{N}_2\text{O}$ , 1.0 g/L  $\text{MgCl} \cdot 6\text{H}_2\text{O}$ , 0.2 g/L  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 0.1 g/L  $\text{FeCl}_2 \cdot 6\text{H}_2\text{O}$ , 1.0 g/L L-Cysteine HCl monohydrate, 0.5  $\mu$ L trace minerals, 0.02 g/L pyridoxine HCl, 0.004 g/L 4-aminobenzoic acid, 0.002 g/L D-biotin, 0.002 g/L vitamin B12 in an anaerobic chamber (COY Laboratory Products, Inc., Grass Lake, MI, USA). In order to remove the residual yeast extract, a second culture was grown overnight by 1% transfer from the first culture. The same procedure was repeated for the third time and the culture obtained after the 3rd batch transfer was used for inoculation of the fermentation bottles. Fermentations were conducted in a 125 mL glass bottle sealed with a butyl rubber stopper [156] in 50 ml volume.

The headspace was purged with nitrogen and the bottles were shaken at 200 rpm at 55 °C for ~72 h.

## **2.8. Enzyme Assays**

### **2.8.1. Preparation of Cell Free Extract**

Cell extract was prepared as described previously [6]. Briefly, 50-ml cultures of *T. saccharolyticum* cells was grown anaerobically in CTFUD medium to OD600 = 0.3-0.6. Collected cells were harvested by centrifugation at  $6000 \times g$  for 15 minutes at RT and stored at  $-80^{\circ}\text{C}$  until assays. The pellet was washed twice anaerobically with a 1 mL deoxygenated buffer (0.1 M Tris-HCl and 0.1 Mm DTT, pH 7.5) just before the assay. The cells were incubated with 3  $\mu\text{l}$  Ready-Lyse Lysozyme (20 KU/ $\mu\text{l}$  to 40 KU/ $\mu\text{l}$ ) for 30 min at room temperature and at the end of the incubation time 2  $\mu\text{l}$  DNase I (25 U/ $\mu\text{l}$ ) solution was added to reduce viscosity. Then the reaction tubes was incubated at room temperature for 20 min. After centrifugation of the the crude lysate at  $12,000 \times g$  for 5 minutes at RT. Obtained supernatant was used as CFE (cell-free extracts) for enzyme assays.

The amount of the total protein was determined by Bradford assay.

### **2.8.2. Alcohol Dehydrogenase (ADH) - Aldehyde Dehydrogenase (ALDH) Assays**

The reactions was performed as described before [157] in 1.2 mL total volume in reduced-volume quartz cuvettes (part number 29MES10; Precision Cells Inc., NY) with a 1.0 cm path length. For ADH (acetaldehyde reduction) and ALDH (acetyl CoA reduction) reactions the mixture of the reaction was prepared as follows; 50 mM Tris-HCl buffer (pH 7.5), 0.3 mM NADPH or NADH, 10 mM acetaldehyde (ADH) or 1 mM acetyl CoA (ALDH), 2 mM  $\text{MgCl}_2$  and 1 mM DTT. Specific activities were evaluated at  $55^{\circ}\text{C}$ , by decrease in absorbance at 340 nm caused by NADPH or NADH oxidation inside an anaerobic chamber. Reactions were monitored by an Agilent 8453 UV-vis spectrophotometer with Peltier controlled temperature (part number 89090A) to maintain assay temperature. Reaction was started with the addition of acetaldehyde or acetyl-CoA. To determine the back ground activity the absorbance of the reaction was measured before

the reaction was started by acetaldehyd addition and was then subtracted from the activity. For all enzyme activities the units are denoted as  $\mu\text{mol of product} \cdot \text{min}^{-1} \cdot (\text{mg of cell extract protein})^{-1}$ . In order to confirm that specific activities was proportional to the amount of extract added, at least two concentrations of cell extract were used for each enzyme assay.

### **2.8.3. Hydrogenase Assays**

Hydrogenase assays were performed anaerobically at 60°C in anaerobic microcuvettes with rubber stoppers (Starna Cells, Atascadero, CA) sealed in an anaerobic chamber (COY Labs, Grass Lake, MI) with an atmosphere of ~89% N<sub>2</sub>, 10% CO<sub>2</sub>, 1% H<sub>2</sub>. Initial rates of benzyl viologen (BV) reduction were measured with a spectrophotometer (Agilent 8453 UV-Vis spectrophotometer with Peltier heat control) at 578 nm ( $\epsilon = 8.65 \text{ mM}^{-1} \text{ cm}^{-1}$ ) [158]. The reaction was prepared in a 1 mL final volume in 50 mM Tris-HCl (Ph 7.5) buffer which also included 1.0 mM BV and 0.2 to 0.06  $\mu\text{g}$  of protein. The reaction was initiated with the addition of 0.02 mmol of hydrogen to the cuvette headspace [30].

### 3. RESULTS

#### 3.1. Deletion of the *hfsA*, *hfsB*, *hfsC* and *hfsD* Genes in *T. saccharolyticum* with *KanR* replacement

*T. saccharolyticum* metabolism consists of three putative hydrogenases; an NAD-dependent [FeFe]-hydrogenase (*hyd*), a probable [Ni-Fe] hydrogenase (*ech*) and a Fe-only-hydrogenase which is assigned as *hfs* (hydrogenase-Fe-S) containing Fe-S clusters and a PAS (Per-Arnt-Sim) domain which can function as a redox sensor in some bacteria [30]. The primary hydrogenase in *T. saccharolyticum* is called *hfs*, and it consists of 4 subunits, *hfsA*, *hfsB*, *hfsC* and *hfsD* (Figure 11). In a study conducted by Saw et. al. [30] the whole *hfs* operon was disrupted with a *KanR* replacement, and it was appeared that removal of only the *hfsABCD* was not sufficient to produce higher ethanol, although the ~ 95% less acetic acid and hydrogen produced in the  $\Delta hfsABCD$  mutant (LL1187) when compared to wild-type. However, after a few generations, it has been reported that there was some increase in ethanol production. When these strains were re-sequenced, some important point mutations were observed in the *hfsB* region.

To better understand the function of the *hfs* operon in *T. saccharolyticum*, each subunit (A, B, C, and D) were deleted by allelic replacement with the kanamycin antibiotic resistance cassette.

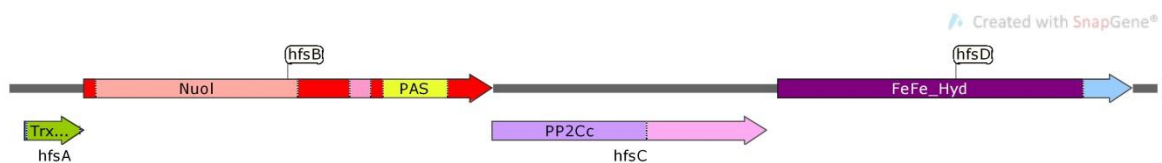


Figure 11. *hfsABCD* operon (4.86 kb) chromosomal organization and domain structure scheme in *T. saccharolyticum*



To inactivate the genes of the *hfs* operon, each assembly fragment for homologous recombination was selected to be approximately 1000 bp up-downstream of the appropriate regions of the *T. saccharolyticum* genome with overlaps as shown in Figure 12 and ~1000 bp of upstream region (5' flanking region) (Figure 12 lines 1, 4, 7 and 10), and ~1000 bp of downstream region (3' flanking region) (Figure 12 lines 3, 6, 9 and 11) were obtained. ~1633 bp of *kan* resistance marker fragment was also amplified with overlaps at both ends (Figure 12 lines 2, 5, 8 and 11). These fragments were assembled via Gibson assembly and resulting replacement cassettes were analyzed on 1% agarose gel after amplification with PCR using 5' flank forward and 3' flank reverse primer pairs (Figure 13). Replacement cassettes were generated as expected band sizes of 3.579 kb, 3.881, 3.738 and 3.620 for *hfsA*, *hfsB*, *hfsC* and *hfsD* deletion, respectively.

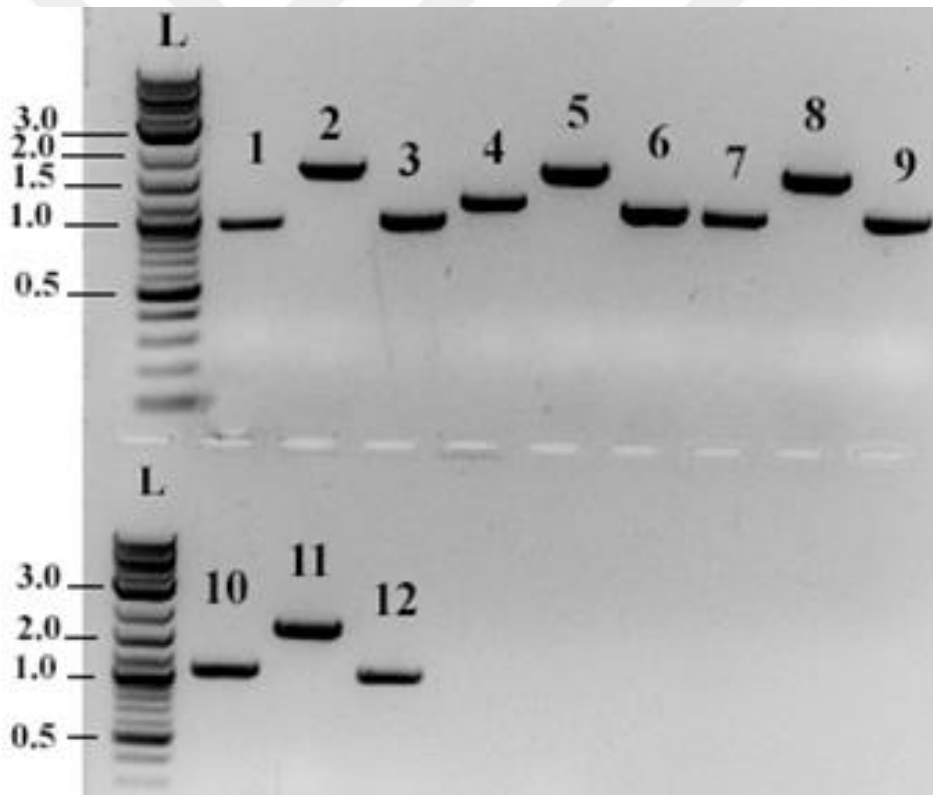


Figure 12. *hfs A, B, C* and *D* deletion fragments generated by PCR, L: DNA Ladder (NEB-N0469S); 1: *hfsA* 3' flank; 2: *hfsA KanR*; 3: *hfsA* 5' flank; 4: *hfsB* 3' flank; 5: *hfsB KanR*; 6: *hfsB* 3' flank; 7: *hfsC* 5' flank; 8: *hfsC KanR*; 9: *hfsC* 3' flank; 10: *hfsD* 5' flank; 11: *hfsD KanR*; 12: *hfsD* 3' flank

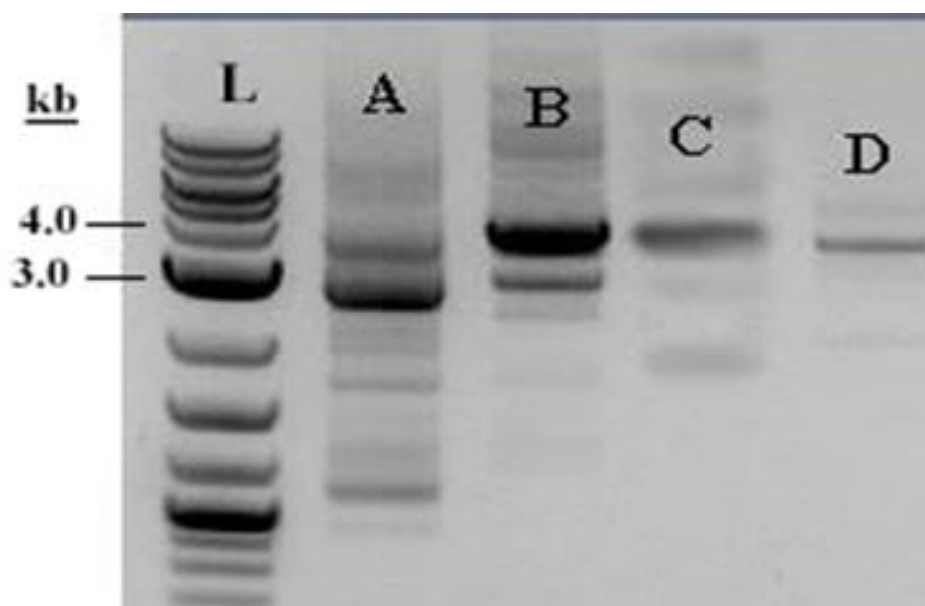


Figure 13. *hfsA*, *B*, *C* and *D* deletion cassettes which were verified and generated by PCR from assembled fragments, L: DNA ladder (2-log quick load purple- NEB); A: *hfsA* replacement cassette (AE2-AE6), B: *hfsB* replacement cassette (AE7-AE11), *hfsC* knockout cassette (AE12-AE16) and D: *hfsD* replacement cassette (AE17-AE21)

After transformation, the resulting mutant *hfs* strains,  $\Delta hfsA::kanR$  (Figure 14),  $\Delta hfsB::kanR$  (Figure 15),  $\Delta hfsC::kanR$  (Figure 16),  $\Delta hfsD::kanR$  (Figure 17) were confirmed to contain approximately 3.6 kb *kan* cassettes in place of the corresponding gene. These new mutants were assigned as LL1267, LL1268, LL1269, and LL1270, respectively.

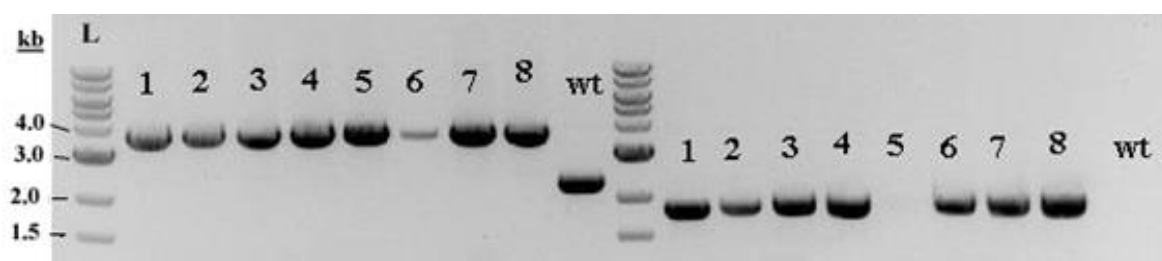


Figure 14. Confirmation of *hfsA* deletion by colony PCR in selected colonies, L: 1 kb DNA ladder (N3232L); 1, 2, 3, 4, 5, 6, 7 and 8 are selected colonies, wt: wild-type, 1.5 2.0 kb bands show amplifications performed with primers targeting the internal region of the *KanR* gene

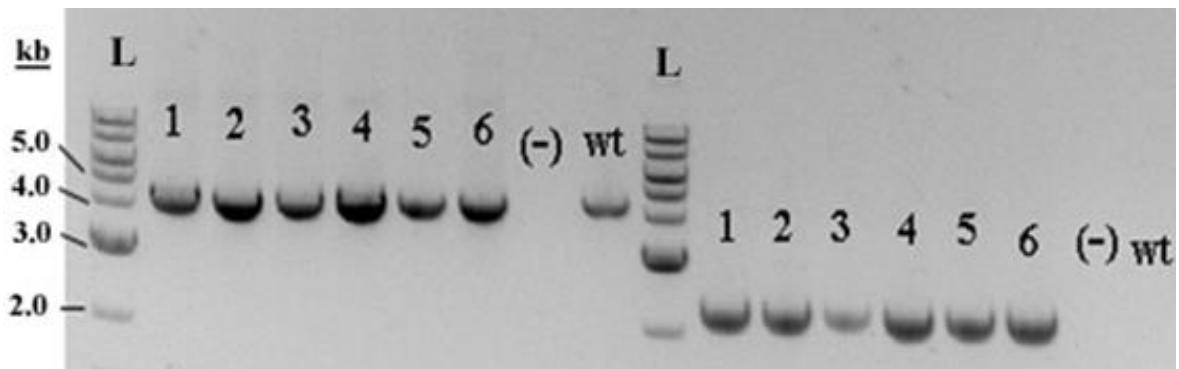


Figure 15. Confirmation of *hfsB* deletion by colony PCR in selected colonies, L: molecular weight marker (1 kb DNA ladder, N3232L), 1, 2, 3, 4, 5, 6, 7 and 8 are selected colonies, wt: wild-type, 1.5-2.0 kb bands show amplifications performed with primers targeting the internal region of the *KanR* gene

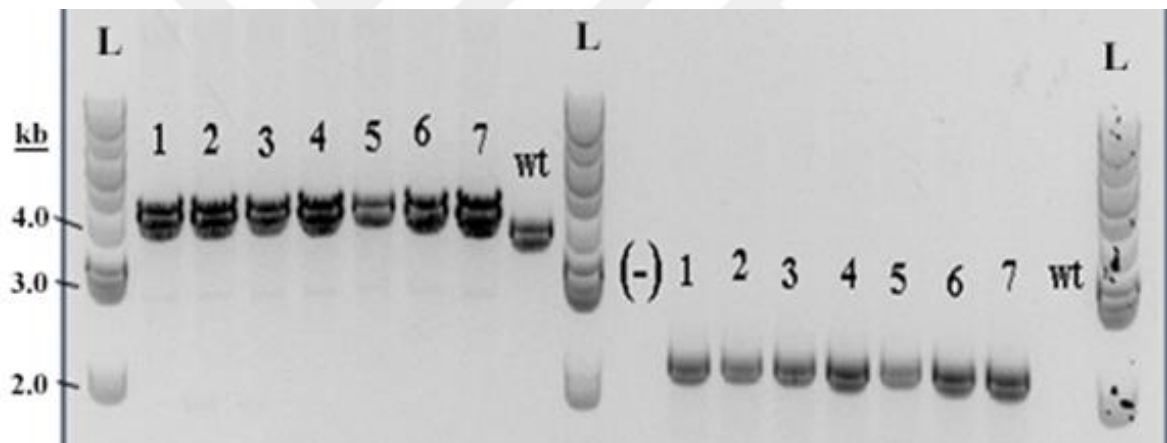


Figure 16. Confirmation of *hfsC* deletion by colony PCR in selected colonies, L: molecular weight marker (1 kb DNA ladder, N3232L); 1, 2, 3, 4, 5, 6, 7 and 8 are selected colonies, wt: wild-type, 1.5-2.0 kb bands show amplifications performed with primers targeting the internal region of the *KanR* gene

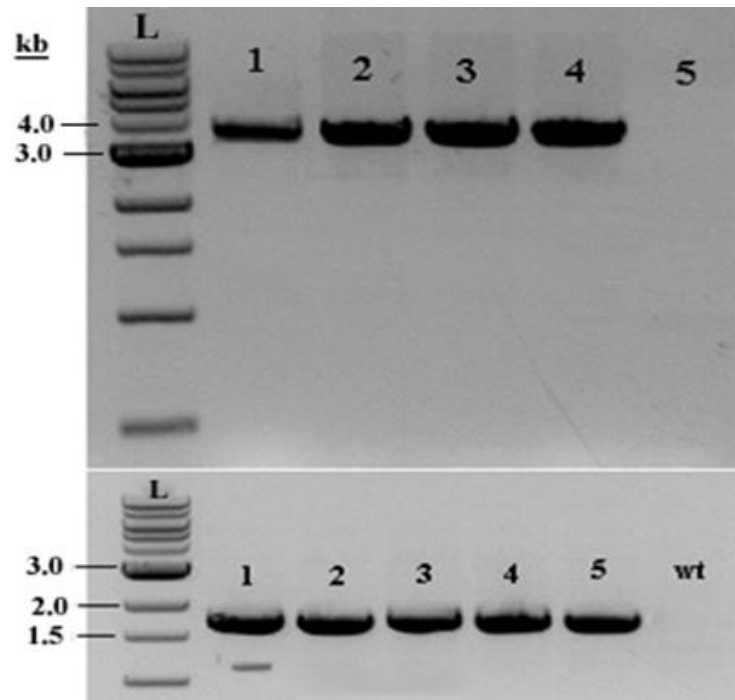


Figure 17. Confirmation of *hfsD* deletion by colony PCR in selected colonies, L: 1 kb DNA ladder (N3232L); 1, 2, 3, 4, 5 are selected colonies, wt: wild-type, 1.5-2.0 kb bands show amplifications performed with primers targeting the internal region of the *KanR* gene

### 3.2. Deletion of *hfsB* Gene in *T. thermosaccharolyticum*, *T. xylanolyticum*, and *T. mathranii*

To determine whether the *hfsB* regulation system is present in other organisms, we deleted *hfsB* in *Thermoanaerobacterium thermosaccharolyticum*, *Thermoanaerobacter mathranii*, and in *Thermoanaerobacterium xylanolyticum*.

PCR amplified 5', 3' and *kan* resistance marker fragments were joined to linear transformation products by Gibson Assembly (Figure 18).

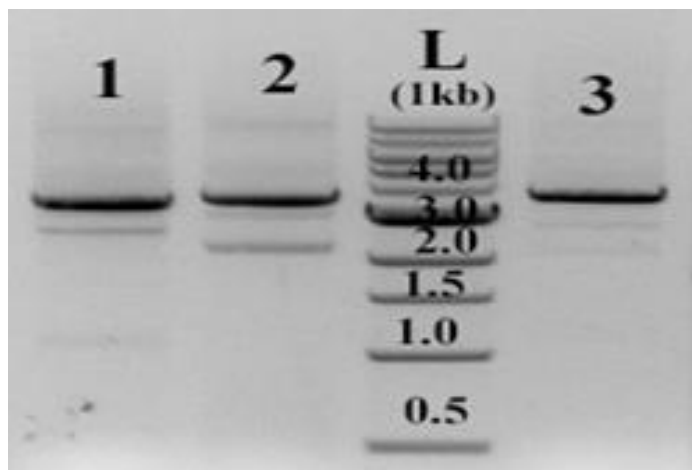


Figure 18. PCR amplified *hfsB* deletion fragments for transformation, L:1 kb DNA ladder (N3232L); 1: for *T. xylanolyticum*; 2: for *T. thermosaccharolyticum*; 3: for *T. mathranii*

Confirmation of gene deletion in *T. xylanolyticum* and in *T. thermosaccharolyticum* by PCR was performed among the colonies selected from the transformation petries, with external and internal primers (Figure 19).

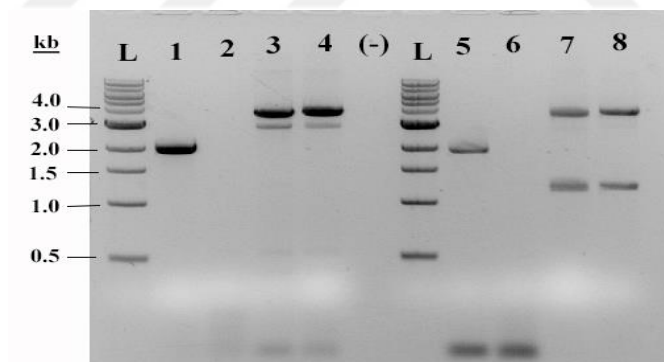


Figure 19. *hfsB* deletion confirmation in *T. xylanolyticum* and *T. Thermosaccharolyticum*, L: 1 kb DNA ladder (N3232L); 1: internal control in LL1347; 2: internal control in LL1301; 3: external control in LL1347; 4: external control in LL1301; 5: internal control in LL1346; 6: internal control in LL1244; 7: external control in LL1346; 8: external control in LL1244

Removal of *hfsB* in *T. mathranii* was confirmed by an internal-external primer pair in selected colonies (Figure 20).

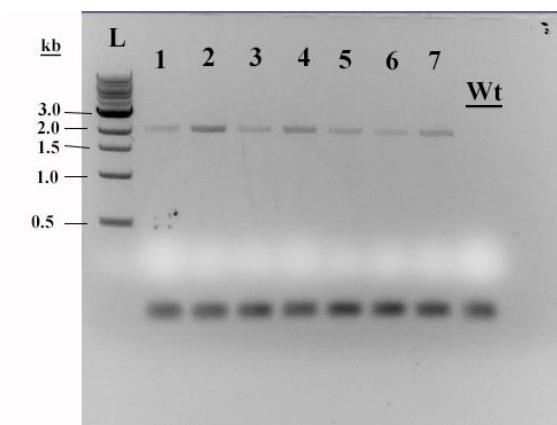


Figure 20. *hfsB* gene deletion internal confirmation in *T. mathranii*;  
L: 1 kb DNA ladder (N3232L); 1, 2, 3, 4, 5, 6, 7 are selected colonies; Wt: LL1258

### 3.3. Removal of *hfsB* in LL345 (*C. thermocellum* $\Delta hpt$ )

Depending on BLAST results it has been shown that *C. thermocellum* genome has a set of *hfsA*, *B*, *C* and *D* genes of *hfs* operon with the same orientation as in *T. saccharolyticum* by the similarity of 59.8%, 53.5%, 59.65% and 33.5%, respectively [16]. Based on the result that was obtained from the *T. saccharolyticum*  $\Delta hfsB::kan$  mutant strain LL1268 it was suggested that deletion of the *hfsB* homologous gene in *C. thermocellum* might give similar fermentation product profile.

In the course of these findings, the probable *hfsB* gene in the *hpt* deletion background strain LL1345 was deleted using plasmid pDG0145 as the backbone (Figure 9). 5' flanking, 3' flanking (1000-1039 bp), internal region (1041 bp) and vector backbone I-II (1976-4974 bp) fragments were generated in accordance with the proper integration for homologous recombination events (Figure 21) and assembly product was transformed into *E. coli*.

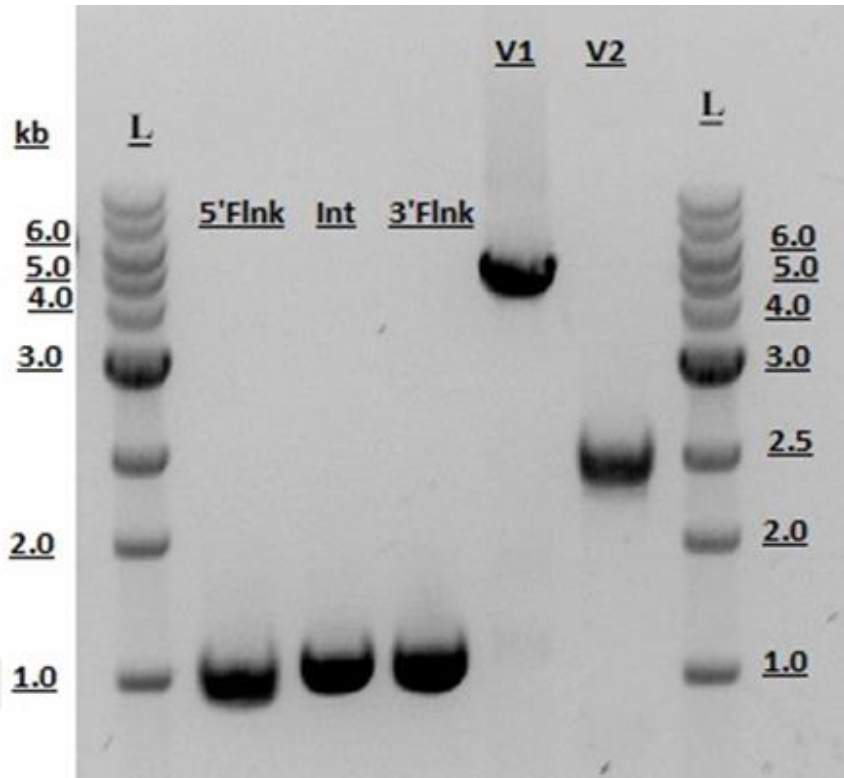


Figure 21. Fragments of *hfsB* gene deletion in *C. thermocellum*, 5' Flnk-Int-3' Flnk: flanking regions for homologous recombination; V1-V2: pDGO145 vector backbones; L: 1 kb DNA ladder (N3232L)

Formed colonies were screened for plasmids that containing 5' flanking-3' flanking regions by colony PCR (Figure 22).

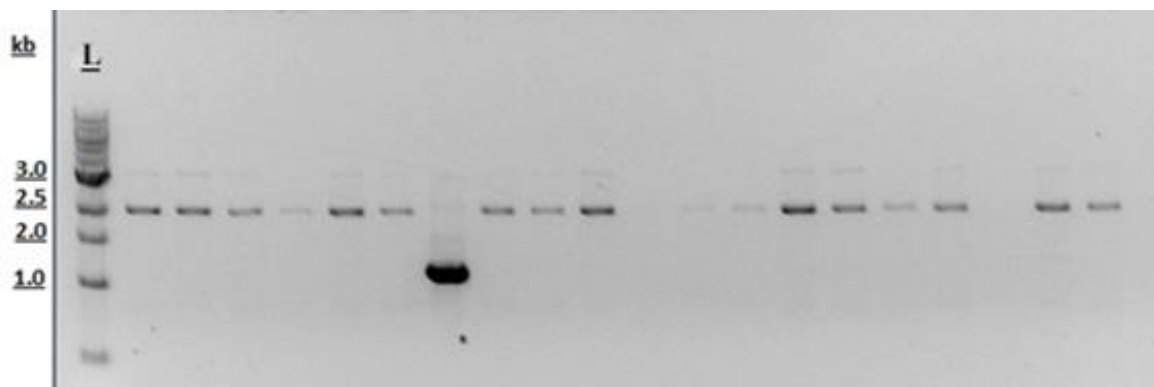


Figure 22. Internal region cloning confirmation in *E. coli* clones with primers AE119-AE120; the band size ~1.0 kb selected as positive clone; L: 1 kb DNA ladder (N3232L)

These colonies were also screened for the presence or absence of the internal region and it was seen that among these 20 colonies only one colony was contained the whole required regions (Figure 23).

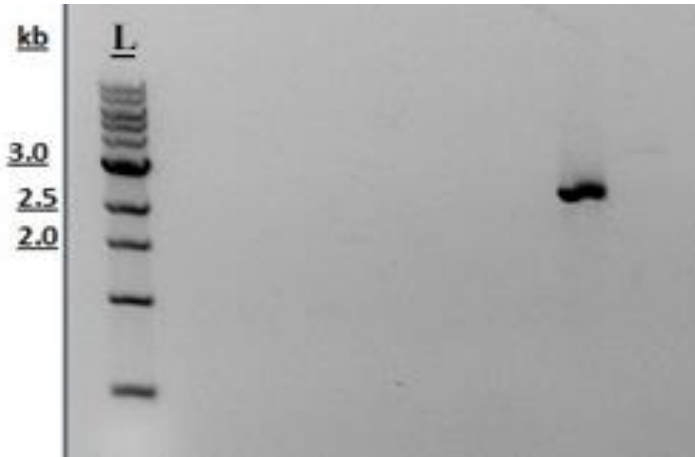


Figure 23. Confirmation of cloning the 5' and 3' flanking regions in *E. coli* clones with primers AE115-AE118; the band size ~2.5 (5' flnk+3' flnk) kb selected as positive clone; L: 1 kb DNA ladder (N3232L)

In this colony, the integrity of the selection markers *cat*, *hpt* and *tdk* were also verified before further steps (Figure 24).



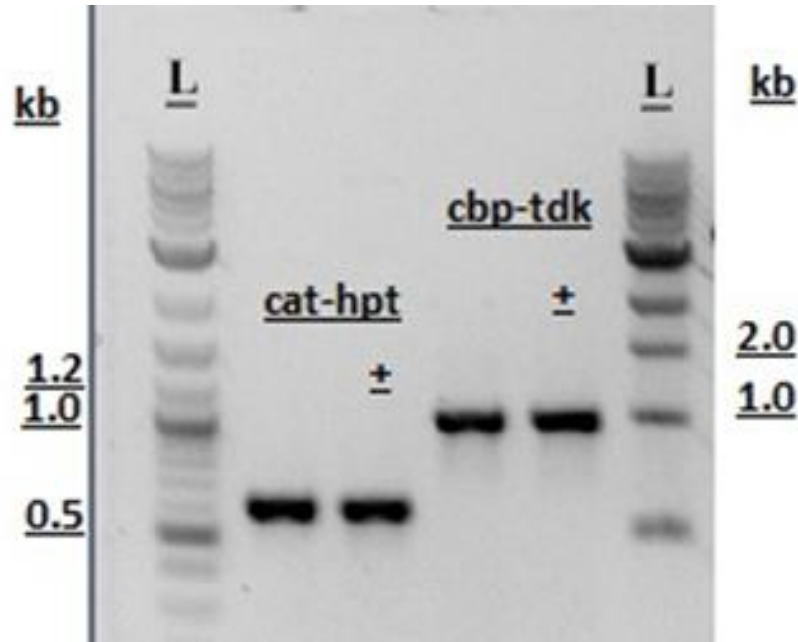


Figure 24. Confirmation of integrity of selection markers in selected *E. coli* clones with primers AE198-AE199 (*cat-hpt*: the band size 616 bp) and AE200-AE201 (*cbp-tdk*: the band size 1029 bp); (+): vector control; L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively

After transformation of the pDGO145-*hfsB* into LL345, selection steps were applied successfully. Among the colonies formed on Tm+FUDR 15 colonies were selected and integrity of selection markers *tdk* (Figure 25) and *cat-hpt* (Figure 26) were verified.

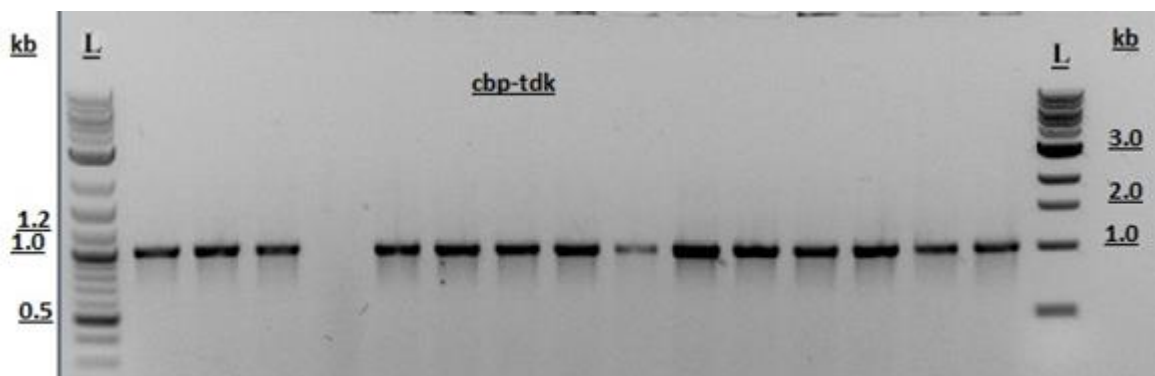


Figure 25. Confirmation of integrity of selection markers *cbp-tdk* in selected *C. thermocellum* clones on Tm-FUDR with the expected size of 1029 bp (primers AE200-AE201); L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively

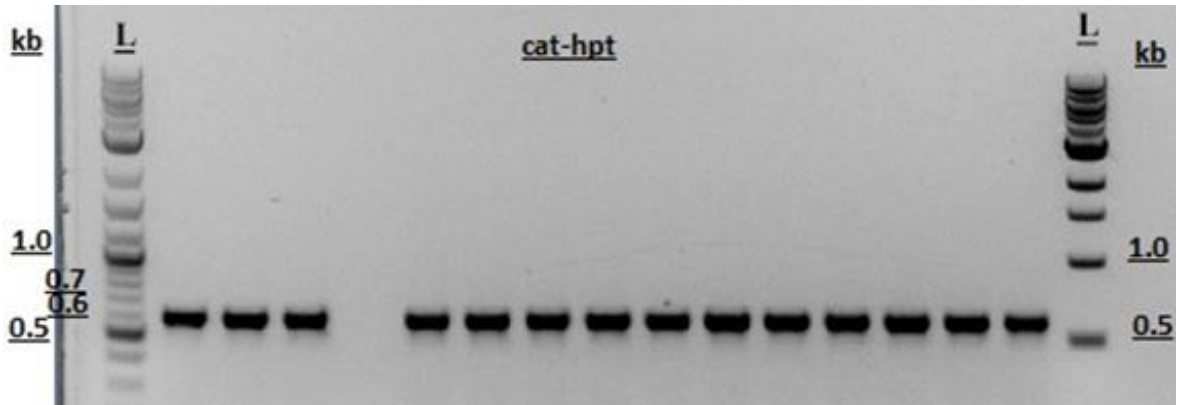


Figure 26. Confirmation of integrity of selection markers *cat-hpt* in selected *C. thermocellum* clones on Tm-FUDR with the expected size of 616 bp (primers AE198-AE199); L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively

At the last step of selection, it was observed that colony formation has occurred as expected on 8AZH containing CTFUD-NY agar and deletion of *hfsB* gene was confirmed by an external primer pair (Figure 27).

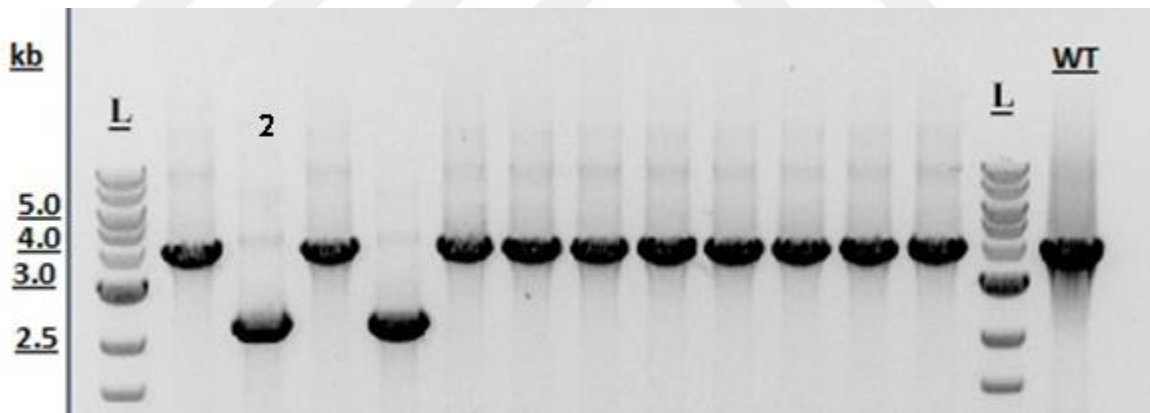


Figure 27. Confirmation of the last step of *hfsB* gene deletion after selection with 8-AZH with the expected band size of 2366 bp (primers AE222-AE223); L: 1 kb DNA ladder (N3232L)

Among 12 colonies, 10 of were false positives and only two were positive. Among these two positive colonies, one of them (colony 2, Figure 27) was selected and plated on CTFUD agar without any selection to ensure purity of the picked colonies. Five of formed colonies were selected from this plate and following the colony PCR by an internal primer pair, no amplification was observed at LL345 size in the selected colonies, while PCR with

the external primer pair was found to have the expected band size of 2366 bp in comparison with the parent strain LL345 (Figure 28).

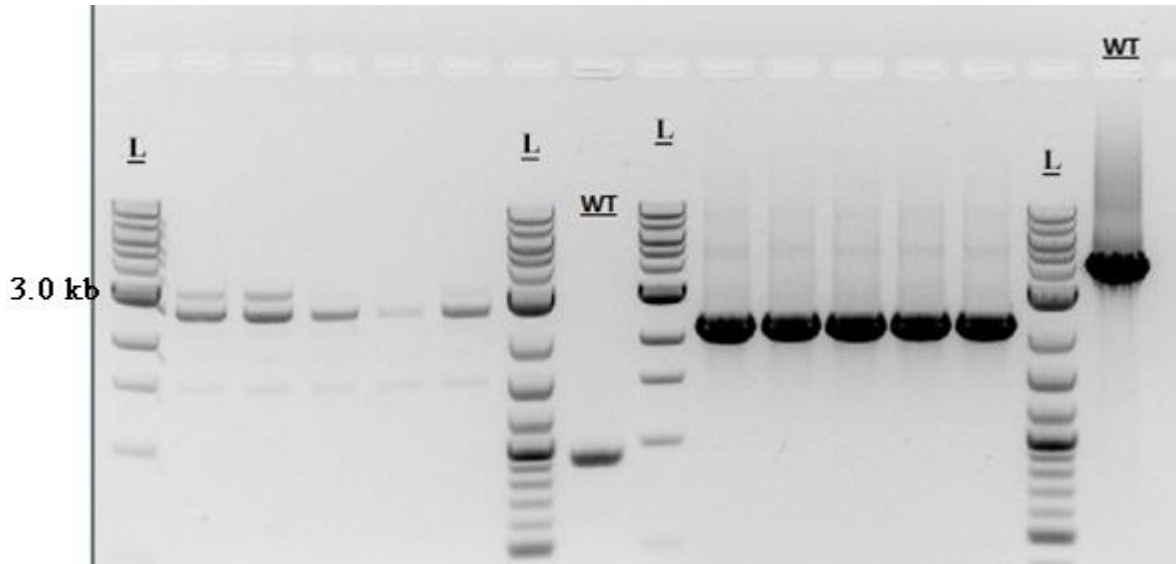


Figure 28. Confirmation of *hfsB* gene deletion in selected *C. thermocellum* clones after last step; L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L), respectively

Complete removal of selection markers were confirmed by PCR in the final mutant strains (Figure 29).

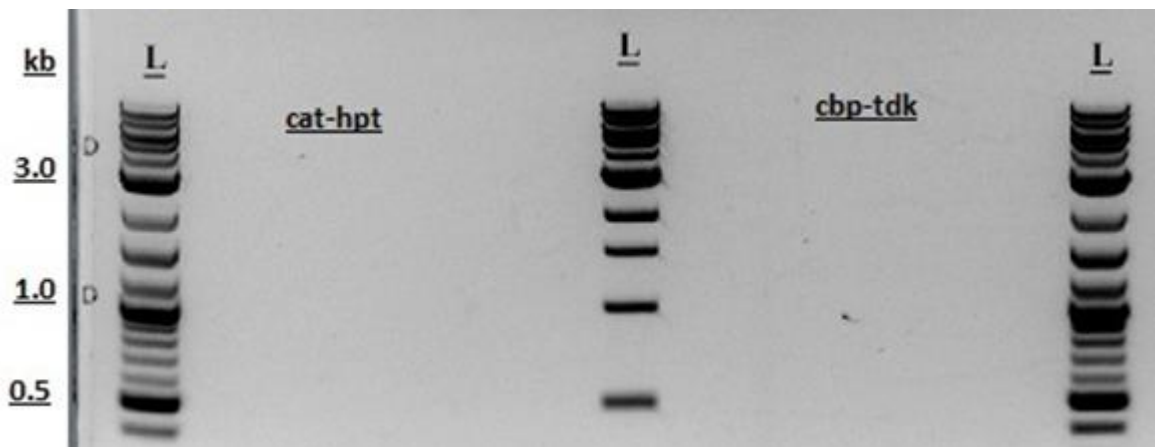


Figure 29. Confirmation of removal of selection markers *cat-hpt* and *cbp-tdk* in selected *C. thermocellum* clones after last step; L: quick load@2-log DNA ladder (N0469S) and 1 kb DNA ladder (N3232L); respectively

### 3.4. Deletion of *hfsAB* and *hfsCD* Clusters in *T. saccharolyticum*

In order to obtain strains in which the *hfsAB* and *hfsCD* genes were inactivated, linear PCR products, including the kanamycin resistance gene from plasmid pMU433 as the selection marker (Figure 30) were transformed into wild-type *T. saccharolyticum* to delete *hfsAB* and *hfsCD* genes.

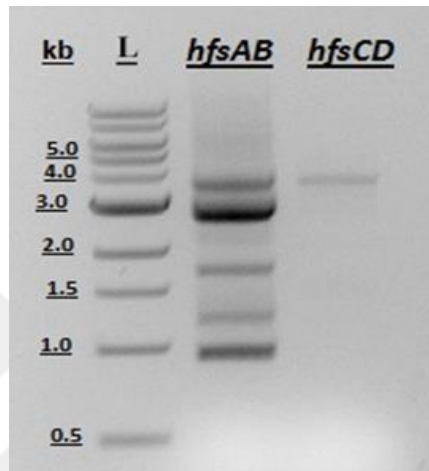


Figure 30. *hfsAB* and *hfsCD* knockout cassette; L: 1 kb DNA ladder (N3232L); *hfsAB* and *hfsCD* knockout cassettes amplified by PCR including kanamycin resistance gene assembled with up and down regions (Up-*KanR*-Down)

In colonies formed after transformation the replacement of the *hfsAB* gene by the *kan* resistance cassette was confirmed using an external primer pair (Figures 31).

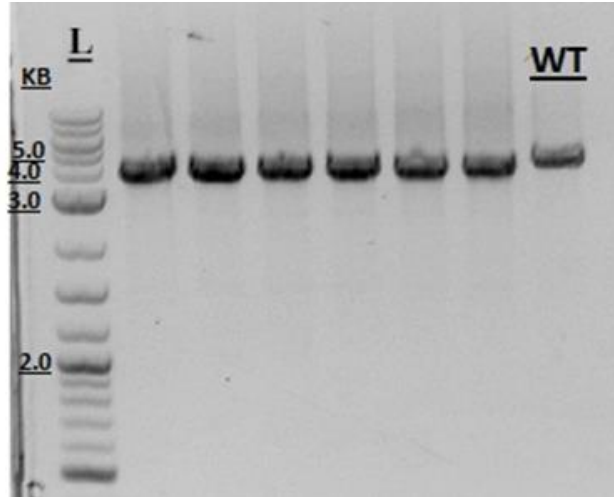


Figure 31.  $\Delta hfsAB$  colony PCR bands with the external primers AE24 and AE27; L: quick load@2-log DNA ladder (N0469S)

Among these, four of the six colonies were given the expected band sizes of 1872 bp with the internal-external primer pair (Figure 32). Clone 1 was selected as the  $\Delta hfsAB::kan$  (LL1349).

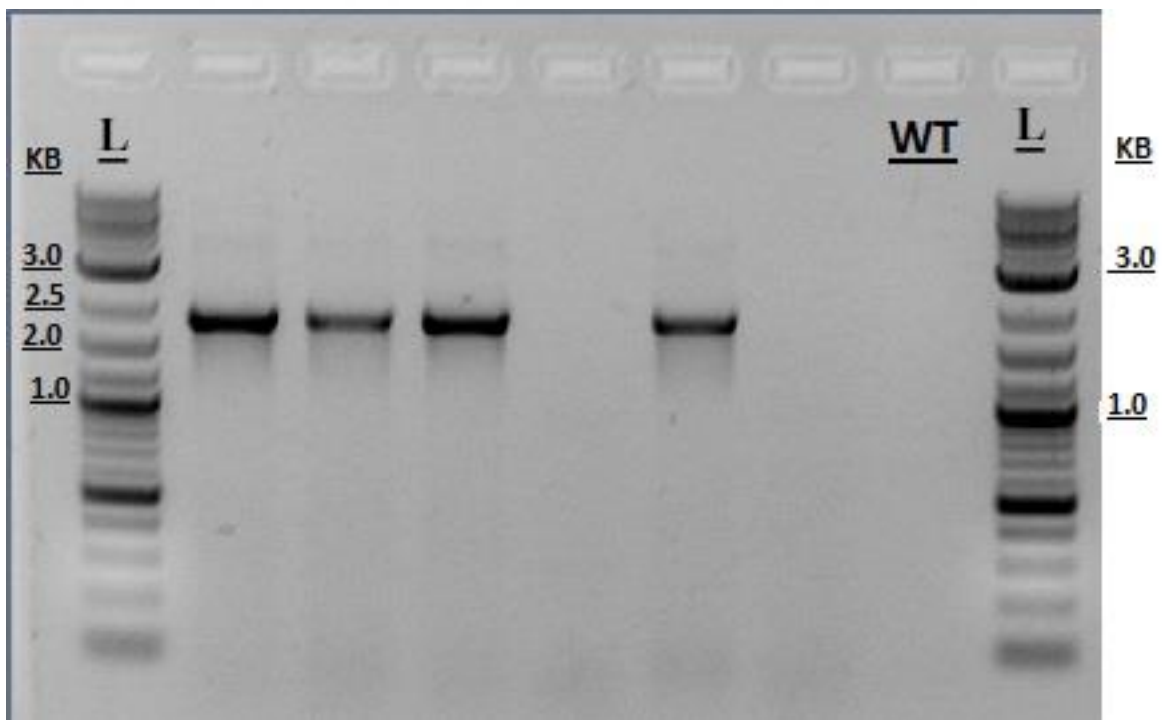


Figure 32.  $\Delta hfsAB$  colony PCR bands with the primers AE24 (external) and AE23 (internal); L: quick load@2-log DNA ladder (N0469S)

Replacement of the *hfsCD* was also verified using an external primer pair in the obtained colonies (Figure 33).

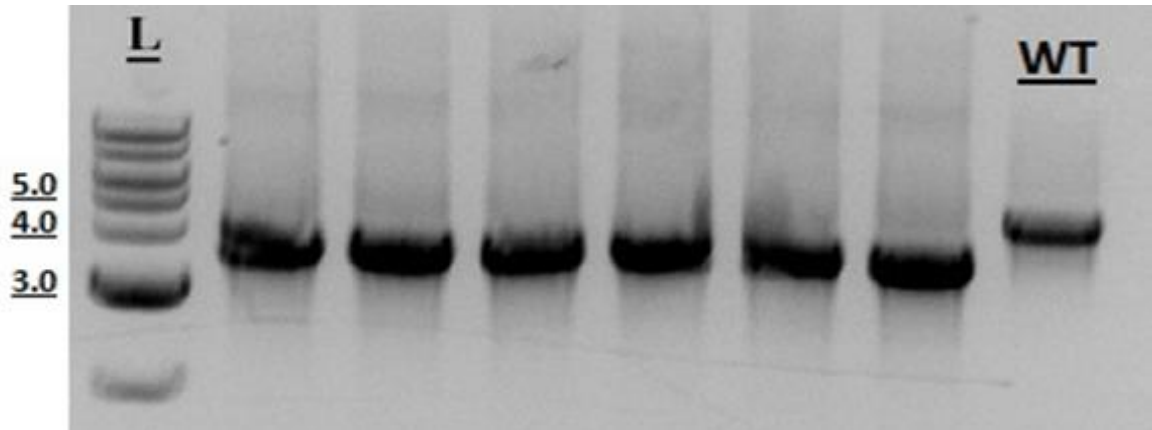


Figure 33.  $\Delta hfsCD$  colony PCR bands with the external primers AE28 and AE31; L: 1 kb DNA ladder (N3232L)

It was observed that in only two colonies products were obtained with the expected sizes from the PCR reaction in which was performed with an internal and an external primer with the expected size of 2036 bp (Figure 34).

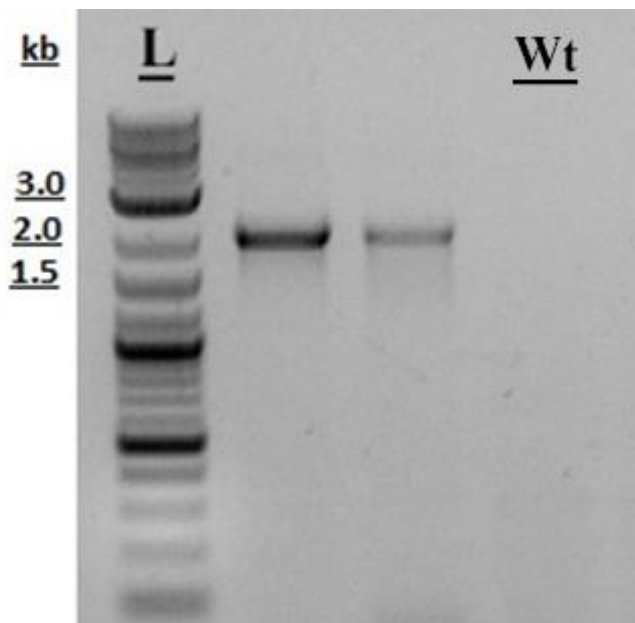


Figure 34.  $\Delta hfsCD$  colony PCR bands with the primers AE28 (external) and AE23 (internal); L: quick load@2-log DNA ladder (N0469S)

Clone 1 was picked as the  $\Delta hfsCD::kan$  (LL1350) strain (Figure 34).

Genomic disruption of the *hfsAB* (LL1349) and *CD* (LL1350) gene clusters has provide an important information about the ethanol increase mechanism. The *hfsAB* (LL1349) mutant strain has shown the same fermentation product disruption with the *hfsB* (LL1268) knockout strain. In the other hand  $\Delta hfsCD$  (LL1350) mutant revealed the same product profile with the  $\Delta hfsC$  (LL1269) and  $\Delta hfsD$  (LL1270) mutant strains. It was suggested that *hfsC* and *hfsD* genes are both essential for high ethanol production in *T. saccharolyticum* when *hfsA* or *B* is absent.

### 3.5. Inactivation of Lactate Dehydrogenase (*ldh*) in LL1267 and LL1268

The elimination of lactic acid as a fermentation product resulted in proportionately increased yields of acetic acid and ethanol, as previously reported by Saw et al. [16]

Depending on this data, it has been suggested that, inactivation of the *ldh* in the  $\Delta hfsA$  and  $\Delta hfsB$  strains could have resulted in some further increase in ethanol production. To test this hypothesis the upstream (1007 bp) and downstream (1069 bp) regions of *ldh* were

generated from *T. saccharolyticum* and assembled with *ErmR* (1337 bp) (Figure 35) from plasmid pMU620 [16].

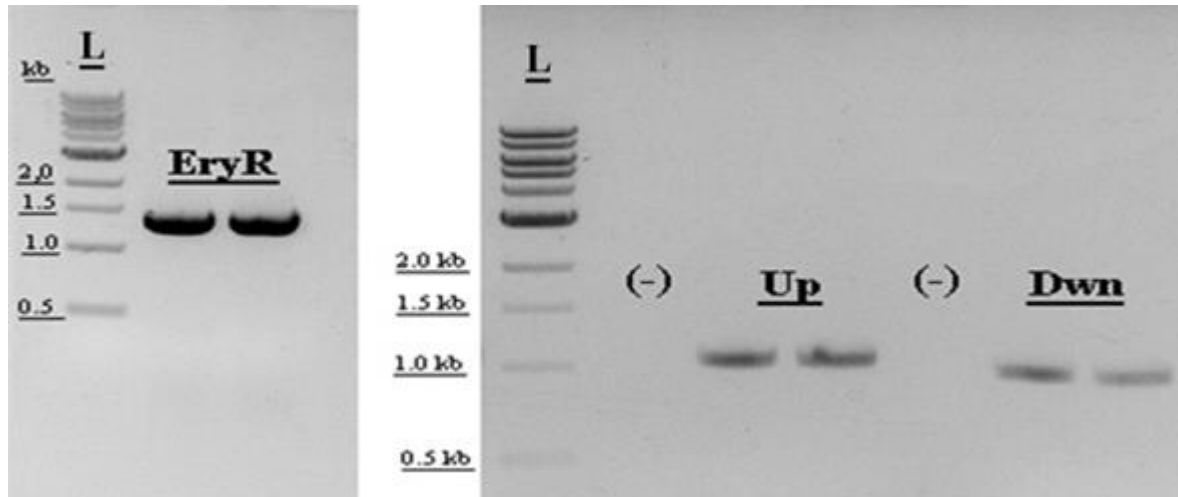


Figure 35. *ldh* knockout fragments generated by PCR, L: 1 kb DNA ladder (N3232L) *EryR*; erythromycin resistance cassette; Up: *ldh* 5' flanking fragment Dwn: *ldh* 3' flanking fragment

About ~4 kb knockout cassette was obtained from the PCR reaction of the assembly product (Figure 36).

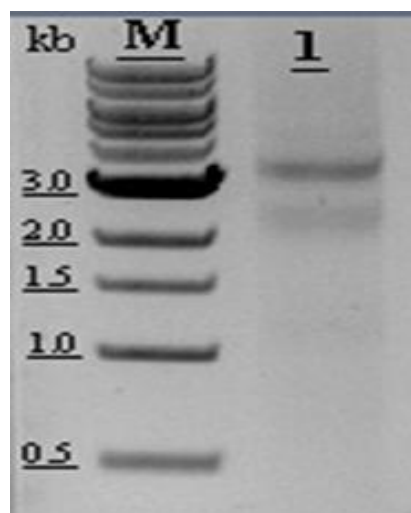


Figure 36. The *ldh* knockout cassette; L: 1 kb DNA ladder (N3232L); line 1: *ldh* knockout cassette including erythromycin resistance gene assembled with *ldh* up and down regions (Up-*EryR*-Dw)



Next after the transformation, 9-10 colonies were subjected to PCR with the external primer pair (Figure 36) and 6 colonies were found harbouring the *erm* resistance cassette in strain LL1267 ( $\Delta hfsA::kan$ ) (Figure 37).

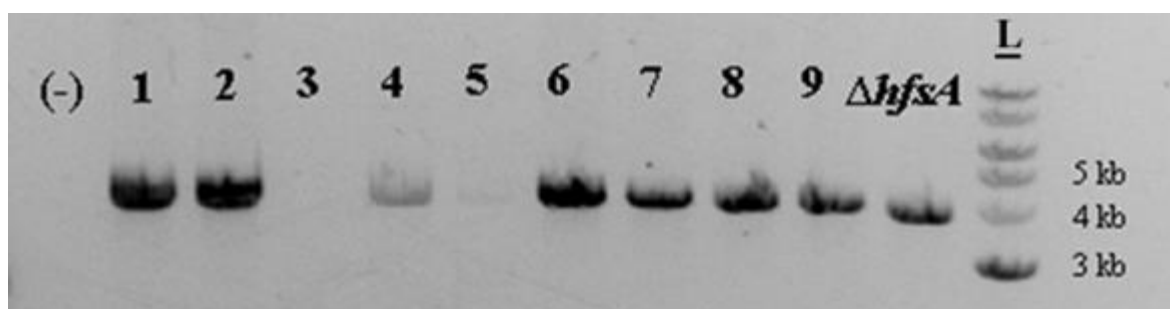


Figure 37. External confirmation of *ldh* knockout in  $\Delta hfsA::kan$  by two external primers via colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); lines 1, 2, 3, 4, 5, 6, 7, 8 and 9 show the ~4.0 kb fragment of *ldh* deletion locus in selected colonies;  $\Delta hfsA$ : *ldh* gene locus amplification fragment in the parent strain by external primers with ~3.9 bp expected size

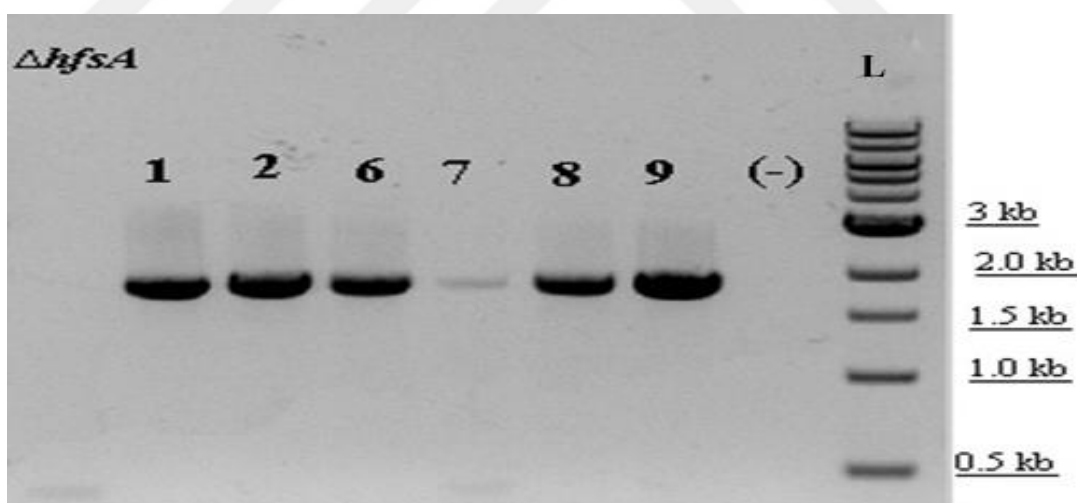


Figure 38. Internal confirmation of *ldh* knockout in  $\Delta hfsA::kan$  by colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); lines 1, 2, 6, 7, 8 and 9 show the ~1.9 kb internal fragment in selected colonies;  $\Delta hfsA$ : negative control for erythromycin cassette

Among the colonies obtained after transformation in LL1268 strain, colony PCR of 8 out of 10 colonies were found to resulted with a proper fragment size with the external

primer pair (Figure 39) while only 6 were observed to containing the *erm* cassette (Figure 40). Colony 2 selected as the target strain for  $\Delta hfsB::kan \Delta ldh::erm$  genotype.

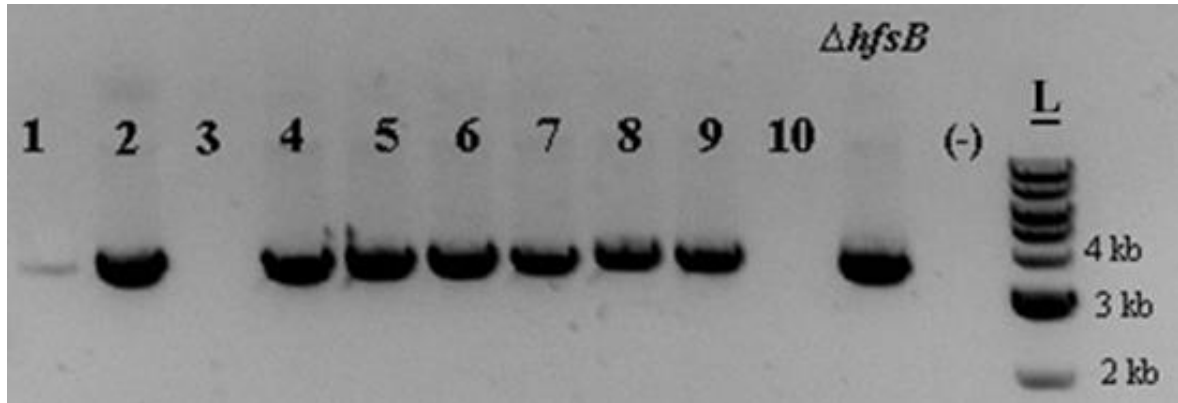


Figure 39. External confirmation of *ldh* knockout in  $\Delta hfsB::kan$  by colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); lines 1, 2, 4, 6, 7, 8, 9 and 10 show the ~4.0 kb external fragment of *ldh* deletion locus in selected colonies;  $\Delta hfsB$ : *ldh* gene locus amplification results in the parent strain by external primers with ~3.9 bp expected size

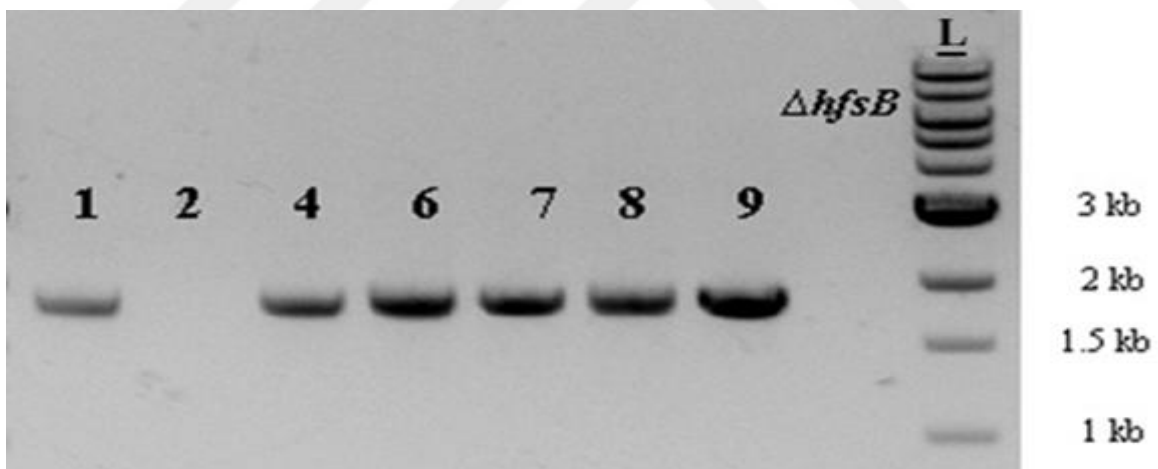


Figure 40. Internal confirmation of *ldh* knockout in  $\Delta hfsB::kan$ , by colony PCR in selected colonies; L: molecular weight marker (1 kb DNA ladder, N3232L); lines 1, 2, 4, 6, 7, 8 and 9 show the ~1.9 kb internal fragment in selected colonies;  $\Delta hfsB$ : negative control for erythromycin cassette

### 3.6. Enzyme Assays

In *T. saccharolyticum*, the conversion of acetyl-Co-A to ethanol is carried out via two reactions, acetyl-CoA reduction to acetaldehyde (ALDH) and acetaldehyde reduction to ethanol (ADH). For these reactions NADH or NADPH can be used as an electron donor.

The enzyme activities were investigated with both electron donors in order to reveal which one is used in reactions, using cell extracts of *T. saccharolyticum* strains. NADH dependent ADH activity of the LL1267 and LL1268 strains was found to be 1.5-2.0 fold higher than the wild-type strain as expected (Figure 41).

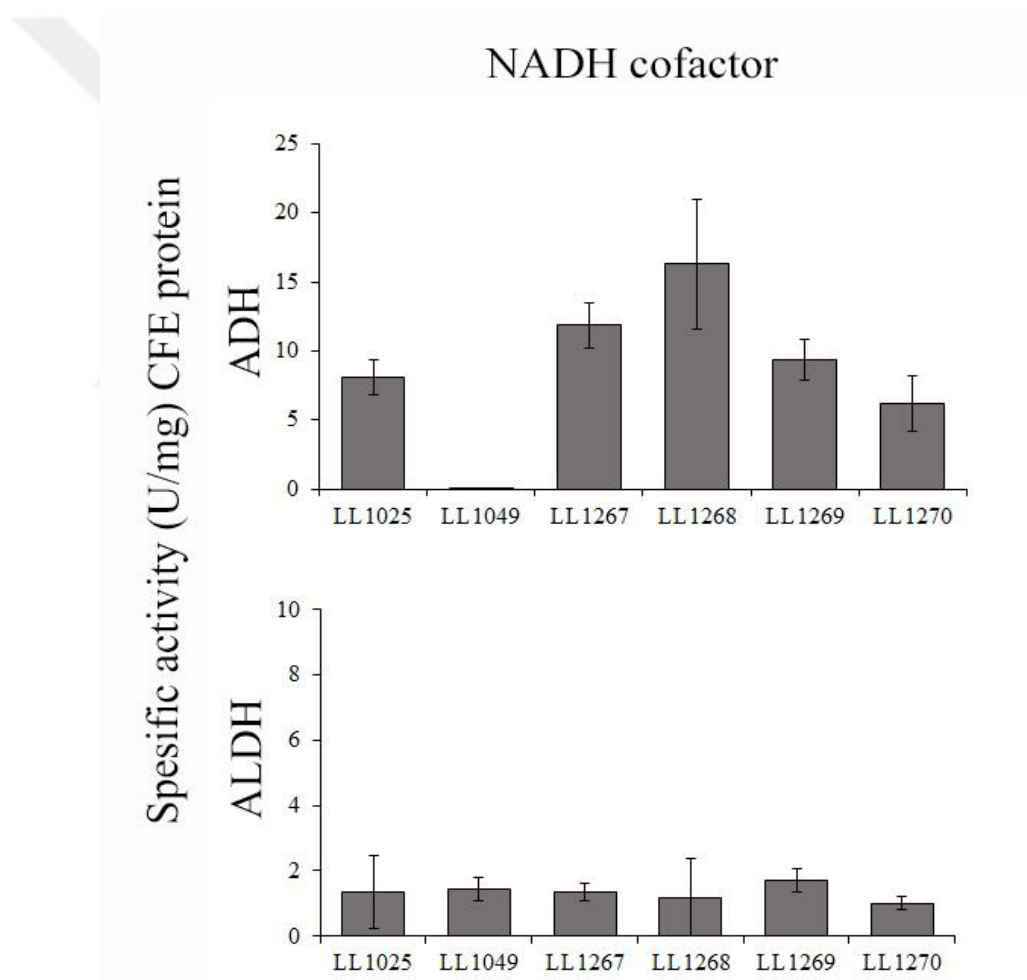


Figure 41. ALDH and ADH activities in *T. saccharolyticum* strains with the NADH cofactor, (SD=1; n ≥ 4)

When NADPH-dependent ADH activity of strains were investigated, it was seen that NADPH was not preferred as a cofactor as much as NADH (Figure 42)

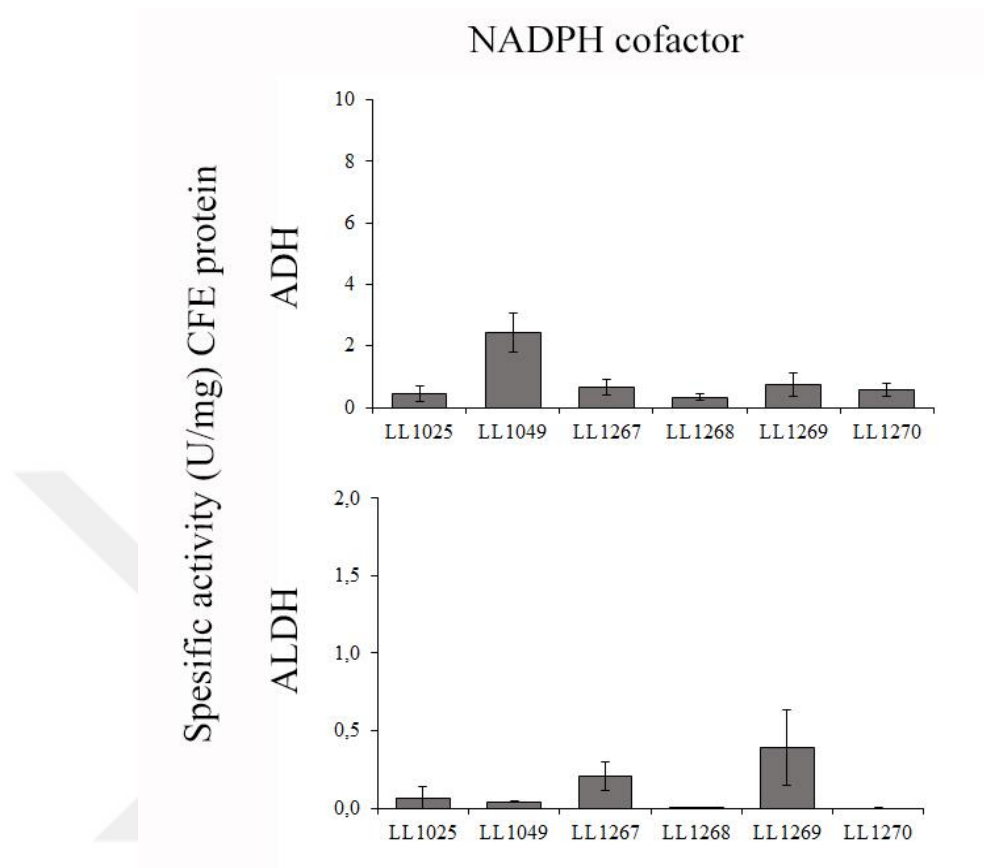


Figure 42. ALDH and ADH activities in *T. saccharolyticum* strains with the NADPH cofactor, (SD=1; n  $\geq$  4)

Benzyl viologen acting as a low potential redox acceptor is an almost universal electron acceptor for and is almost universal electron acceptor for hydrogenases able to interact with hydrogenases that have either NAD(P)H or ferredoxin as a natural substrate [30, 121]. Depending on the conserved domain analysis of the *hfsB* and *hfsD*, it was suggested that these two protein may responsible from the hydrogenase activity in the *hfsABCD* operon. It was found that, the measured BV reduction was increased in LL1268, in contrast to LL1187, when compared to the wild-type strain. Cell extract of the strain  $\Delta hfsD$  was showed lower BV reduction when compared to  $\Delta hfsB$  and wild-type strains (Figure 43)

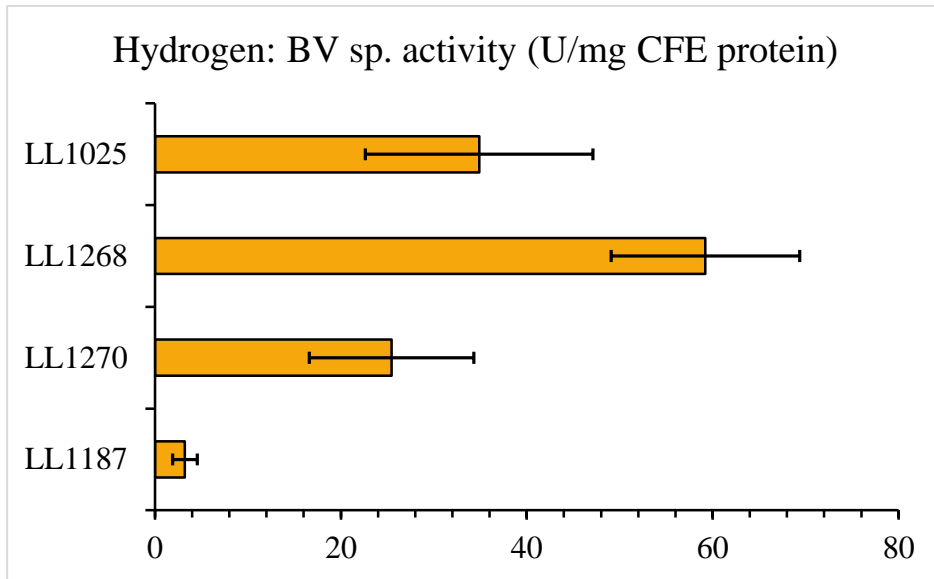


Figure 43. BV reduction in *T. saccharolyticum* strains to test hydrogenase activity, (SD=1; n ≥ 3)

### 3.7. RNAseq and Proteomics Data Analysis

To reveal the change in enzyme activity, transcript levels were measured in different *T. saccharolyticum* strains including the wild type (wt *Tsaccharolyticum*), high-ethanol-producing (LL1049), *hfsABCD* deletion (LL1187) and individual *hfsA*, *B*, *C* and *D* deletions (strains LL1267, LL1268, LL1269 and LL1270) (Figure 44). Also, some preliminary proteomic data were collected. Although this analysis revealed a number of interesting candidates, it was focused on four genes; *adhE*, *adhA*, *rex* and *Tsac\_0415* (Figure 44). The *adhE*, *adhA* and *rex* were selected because they have been studies reported that these genes have play roles in ethanol production in *T. sachharolyticum* [158, 159]. The last was selected because it is adjacent to *adhE*. In order to see if disruption with the kan marker had any polar effects, expression levels of *hfsA*, *B*, *C* and *D* were also investigated.

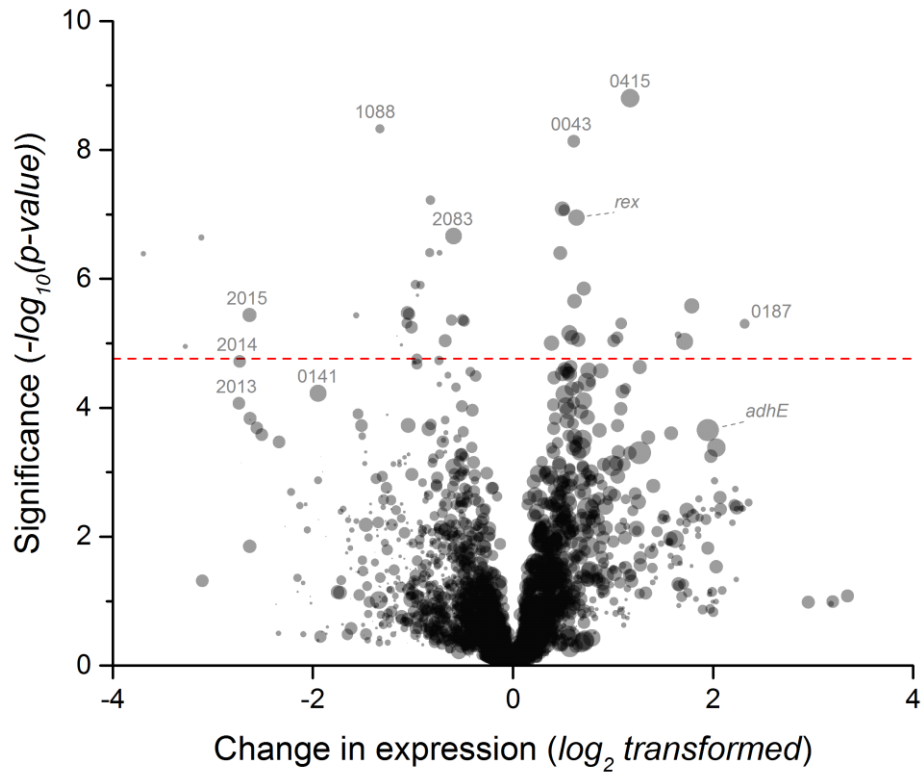


Figure 44. Gene expression analysis of high (LL1267, LL1268 and LL1049) and wild-type level (wt *T. saccharolyticum*, LL1269, LL1270 and LL1187) ethanol-producing strains, the plot is showing difference in gene expression, genes with high expression in the high ethanol producing strains are plotted on the right side, the vertical axis is representing the significance of the change

As a result, it was shown that disruption of *hfsA* reduced expression of downstream genes (*hfsB*, *C*, and *D*) by 2 to 6-fold. Disruption of *hfsB* was not revealed any polar effects. Disruption of *hfsC* slightly down-regulated *hfsD* expression. Disruption of *hfsD* increased *hfsA* and *hfsC* expression but was not changed *hfsB* expression. Depending on these results, it was suggested that insertion of the kan marker did not seem to have a persistence effect on downstream genes. Expression of *Tsac\_0415*, *adhE* and *rex* were all correlated with increased ethanol production (Figure 45).

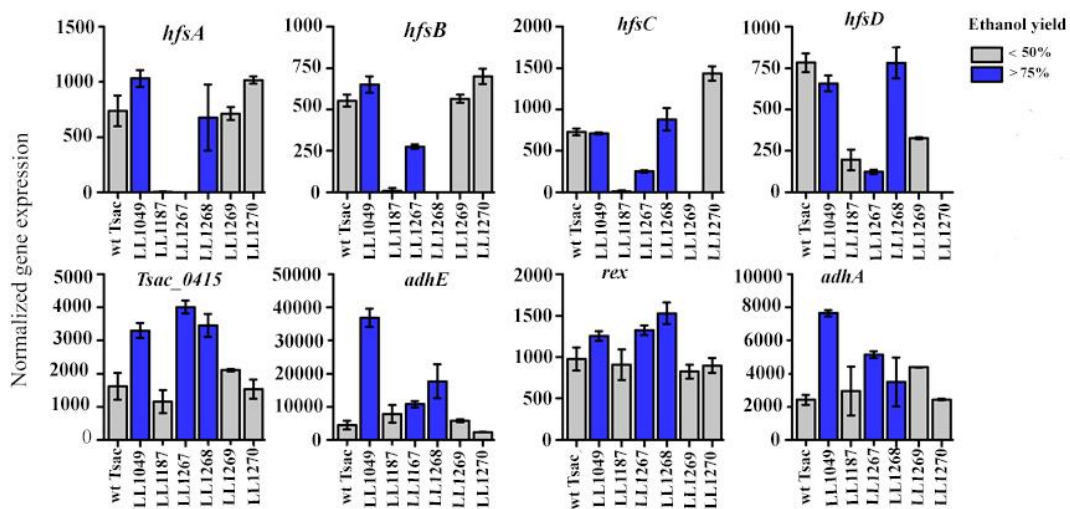


Figure 45. *hfsA*, *B*, *C* and *D* genes expression analysis of the mutant strains, the *hfsA* deletion downregulates *hfsB*, which may explain its phenotype, the other *hfs* deletions have almost no effect on downstream gene expression, upregulation of *adhE* is observed by RT-qPCR, which agrees with RNAseq data

According to the data obtained from proteomics analysis; *AdhE* (Tsac\_0416) still seems to be strongly correlated with increased ethanol production (Figure 38.). The data which acquired from the RNA-seq were revealed that the *hfsA* deletion leads to downregulation of the *hfsB*, *hfsC* and *hfsD* genes expression in the mutant strains. The bifunctional alcohol dehydrogenase gene (*adhE*) found to be upregulated in  $\Delta hfsA$  and  $\Delta hfsB$  strains, as also verified by the proteomics data (Figure 46).

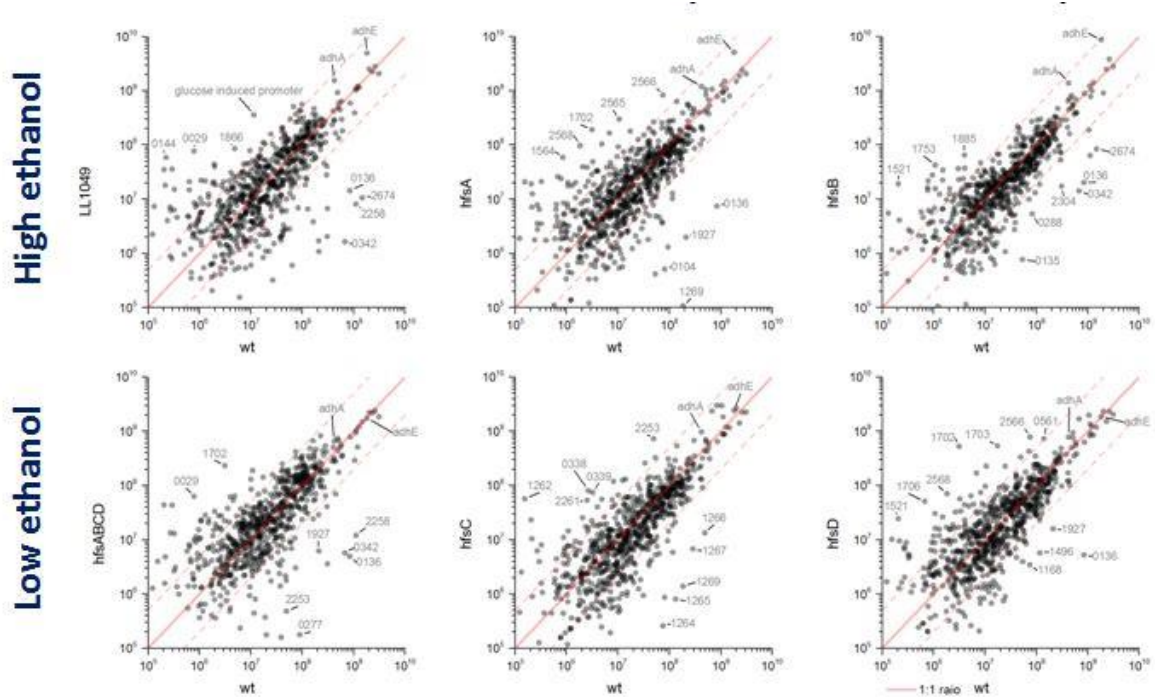


Figure 46. Plot graph of protein abundance vs wt (LC/MS-MS data), genes with 5-fold expression difference compared to wt are indicated with gene name or locus number, *adhE* is consistently over-expressed in high ethanol producing strains

### 3.8. RT-qPCR Expression Analysis

It has been known that the bifunctional ADH/ALDH protein *AdhE* essential for ethanol formation in anaerobic alcohol producers [157, 159]. Since the LL1267 and LL1268 strains produced  $\sim \geq 75$  % ethanol of the theoretical maximum and based on the data that obtained from the proteomics and RNA-seq analysis, we have decided to investigate the real-time expression of the *adhE* gene in the mutant strains. The RNAseq results were confirmed by RT-qPCR for *adhE* and *adhA* (Figure 47).



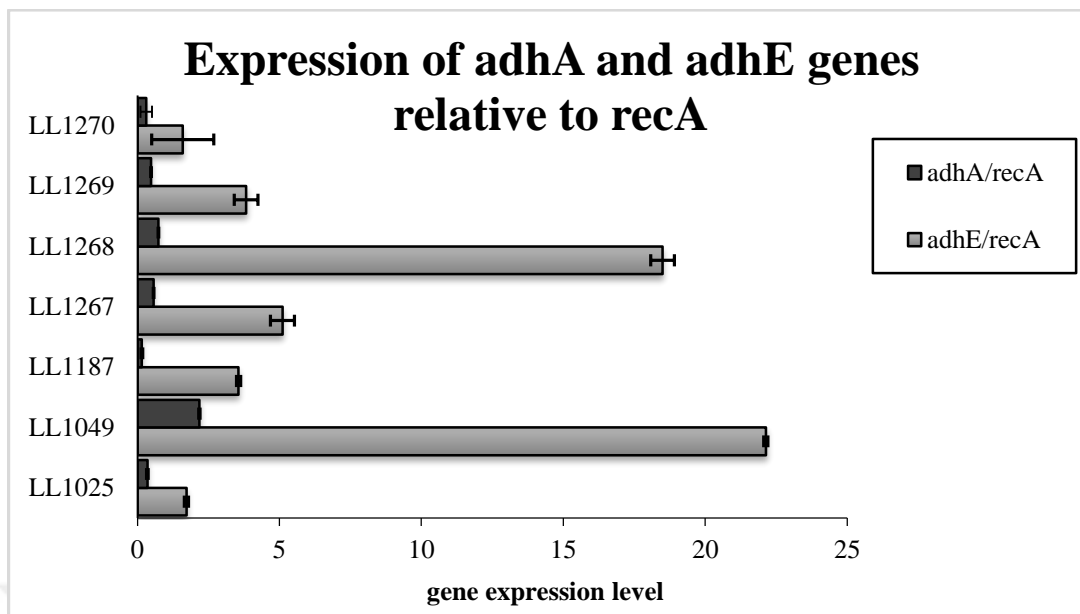


Figure 47. Results of RT-qPCR analysis of *adhA* and *adhE* genes in *T. saccharolyticum* strains normalized to *recA*, (SD=1; n=2)

### 3.9. Whole Genome Resequencing Analysis

In order to verify the complete gene deletions, to control the possible contamination and investigate the possible mutations that may accompany the effect of gene deletions, the whole genome of the mutant strains were resequenced. As a result it was confirmed that all strains that were sequenced are correct and no contamination was found. Secondary mutations in genes known to be related to fermentation product production are noted in Appendix 1. All mutations are listed in Appendix 1. (strains of *T. saccharolyticum*).

Sequence analysis *T. saccharolyticum* has been engineered for ethanol production by deleting pathways for lactate and acetate production [159, 161]. In those strains, ethanol production does not increase immediately after the disruption of the *ldh* and *pta* genes, but does do within a few generations. In order to understand this situation, genomes of *T. saccharolyticum* strains with disruptions of *ldh* and *pta-ack* that exhibited both low and high ethanol yields resequenced (Appendix 1.). In these strains, it was seen that mutations in the *hfsB* gene were related with increased ethanol production. Previously Saw et al. [30] had disrupted the entire *hfsABCD* operon, and had not seen a substantial increase in ethanol production. To figure out this apparent conflict, and to better understand the function of the *hfs* operon in *T. saccharolyticum*, each subunit (*A*, *B*, *C* and *D*) was disrupted one-by-one, using allelic replacement with the kanamycin antibiotic resistance marker (*KanR*).

The analysis results confirmed the clear deletions. When the genome sequence was investigated for SNPs, and/or for indels, it was seen some missense mutation that may not have a substantial effect on the particular gene activity or genes involved in energy metabolism of the related strains.

### **3.10. Fermentation Product Distribution of *T. saccharolyticum* Strains and Analysis**

Following the strain confirmation the effect of the mutations on fermentation end-product distribution in several *T. saccharolyticum* strains were measured (Figure 48). It was seen that strains with disruptions of the *A* and/or *B* subunits (Figure 48, strains LL1267 and LL1268) produced  $\geq 3$  moles of ethanol per mole of cellobiose consumed, whereas strains with disruptions of *C* and/or *D* subunits (Figure 48, strains LL1269 and LL1270) produced ethanol about of  $\geq 2$  molar yield, which equivalent to the wild type strain (wt *T. saccharolyticum*).

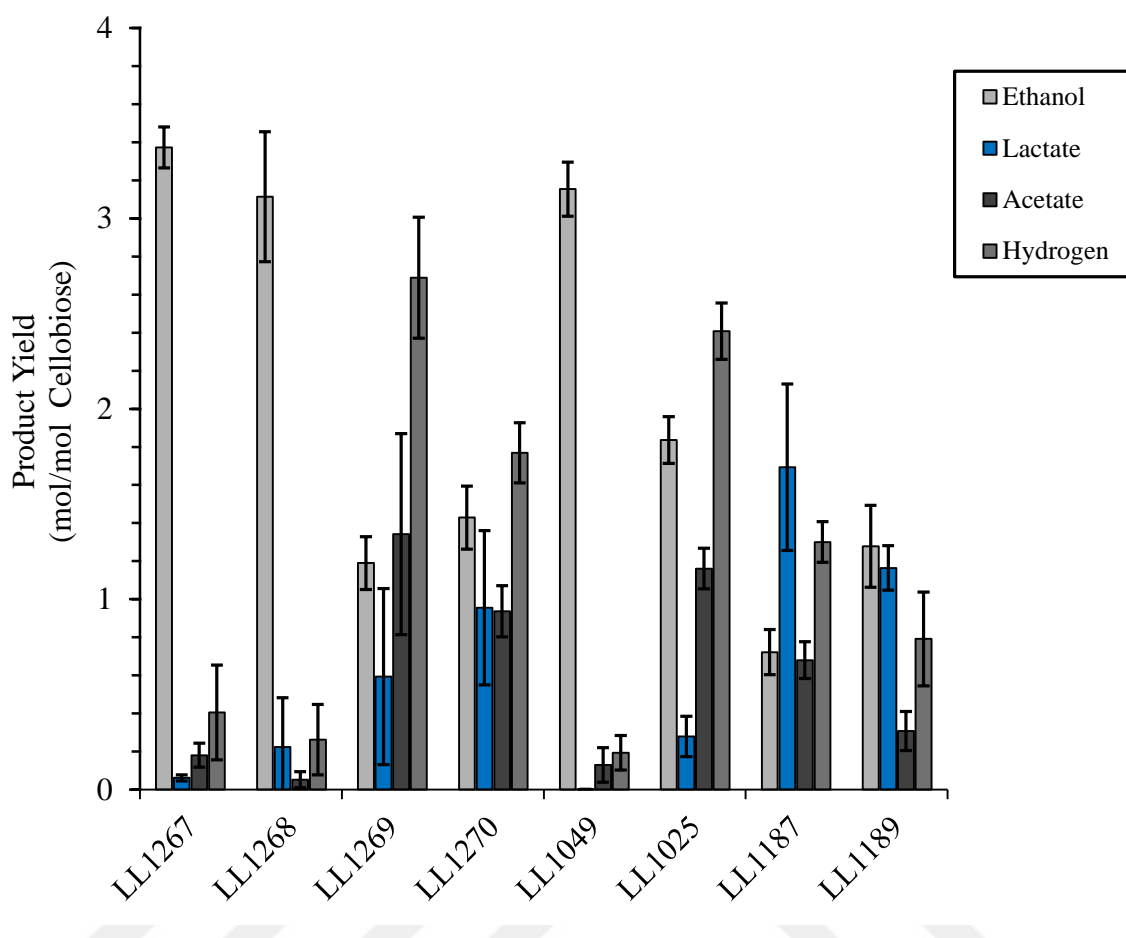


Figure 48. Fermentation product distribution of various *T. saccharolyticum* strains from 5 g/L cellobiose, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n ≥ 2)

Deletion of *hfsA* and *hfsB* in the wild-type strain (Strain ID LL1267 and LL1268, respectively) lead an increase on ethanol production. Also H<sub>2</sub> production have significantly decreased in  $\Delta hfsA$  (70%) and in  $\Delta hfsB$  (90%), which is highly correlated with acetate production. In conclusion deletion of subunit *A* or subunit *B*, but not deletion of subunit *C* or subunit *D*, results in an increase in ethanol yield of  $\geq 75\%$  (Figure 49).

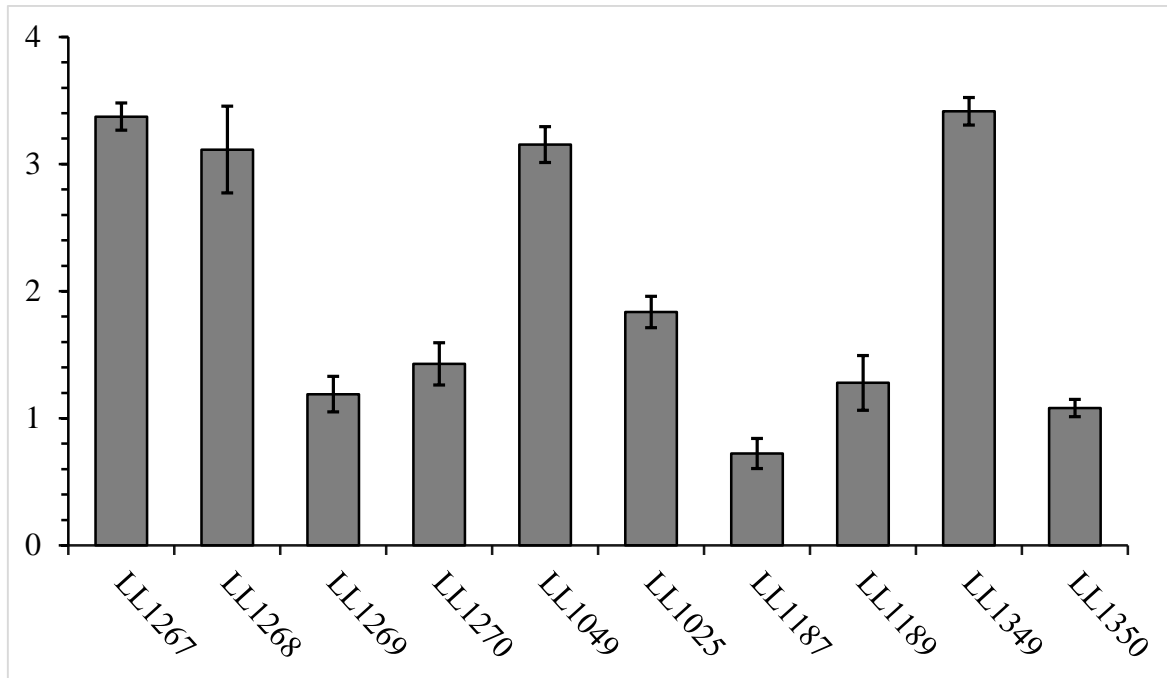


Figure 49. Ethanol yield of several *T. saccharolyticum* strains grew on 5 g/ L cellobiose, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n ≥ 2)

The double gene deletion strains LL1349 ( $\Delta hfsAB::kan$ ) and LL1350 ( $\Delta hfsCD::kan$ ) was showed that the *hfsCD* gene complex must be intact for high ethanol production (Figure 50)

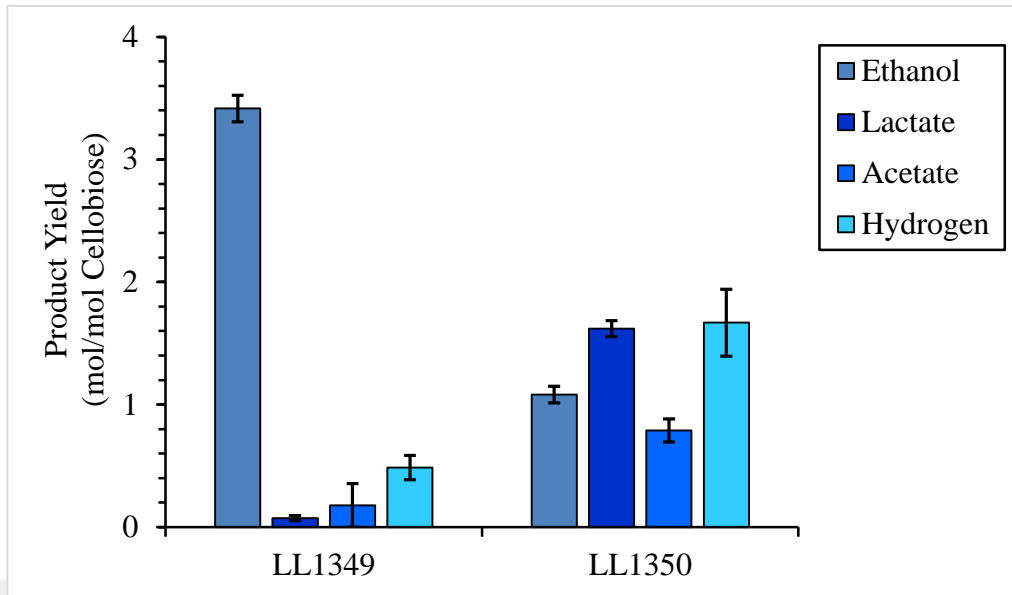


Figure 50. Fermentation product yield of LL1349 and LL1350, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n=3)

The double gene knockout mutant strains  $\Delta hfsA::kan\Delta ldh::erm$  and  $\Delta hfsB::kan\Delta ldh::erm$ , although the erythromycin resistant and the complete loss of *ldh* gene based on the PCR analysis, did not show any significant difference when compared to parent strains  $\Delta hfsA$  and  $\Delta hfsB$  in minimal medium MTC-6 containing 5 g/L cellobiose (Figure 51).

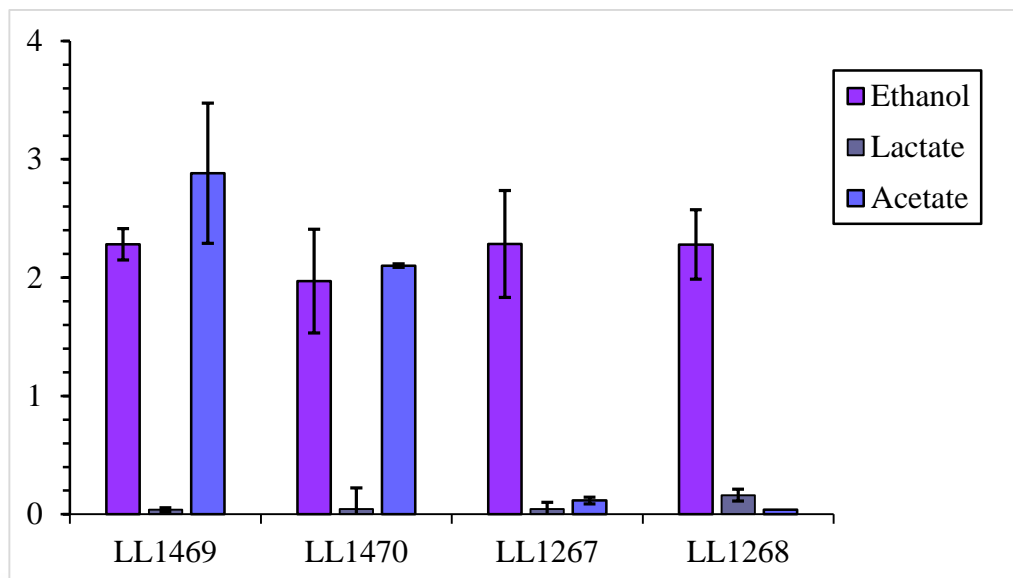


Figure 51. Fermentation product yield of LL1349 and LL1350, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n  $\geq$  2)

### 3.11. Analysis of *T. thermosaccharolyticum*, *T. xylanolyticum* and *T. mathranii* Fermentation Product Distribution

In order to understand whether *hfsB* has similar regulation effect in some other bacteria that close to *T. saccharolyticum* containing a homologous of *hfs* operon, *hfsB* in *T. thermosaccharolyticum*, *T. xylanolyticum* and *T. mathranii* were also disrupted.

It has been observed that the degradation of *hfsB* has significant effects on ethanol production in *T. thermosaccharolyticum* and *T. xylanolyticum* containing the *hfs* operon in the same orientation as the *hfs* operon in *T. Saccharolyticum*. Fermentation analysis results showed that the in these two bacteria, strains with *hfsB* disruption was produced ethanol at a yield of  $\geq 3$  moles of ethanol per mole of cellobiose consumed (Figure 52).

In *T. mathranii*, which has a different *hfs* operon orientation than *T. saccharolyticum*, the ethanol production was remained at a molar yield of about  $<1$  although ethanol production has increased about two fold as compared to wild type (Figure 52)

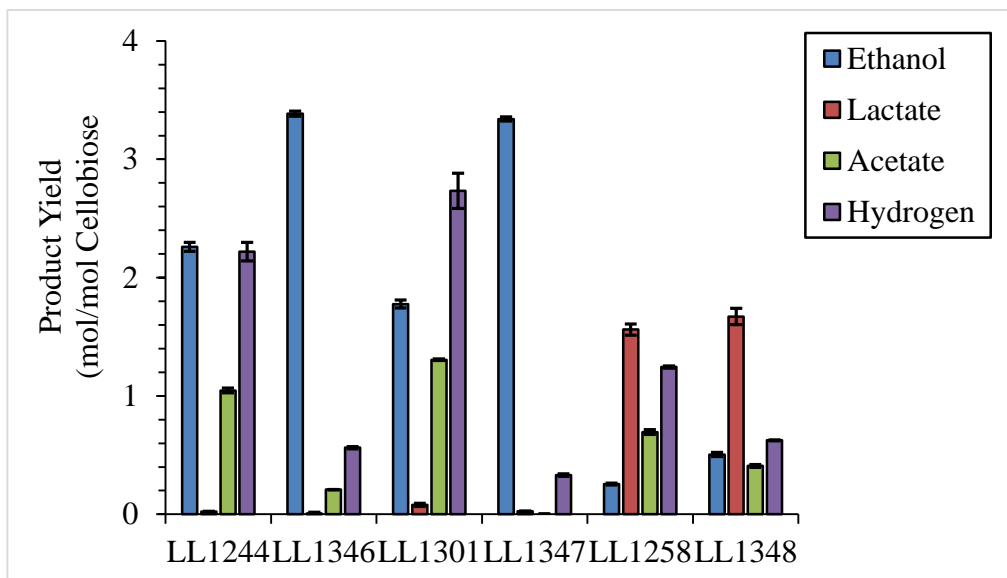


Figure 52. Fermentation product yield of LL1349 and LL1350, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n=3)

### 3.12. Heterologous Expression of *hfsABCD\** Gene in *Clostridium thermocellum* Under Three Different Promoter

In a recent study it has been shown that the *hfs* hydrogenase is responsible for hydrogen production in *T. saccharolyticum*. Deletion of the *hfs* operon resulted in a decrease in hydrogen and acetate, but no change in ethanol production.

Introduction of two point mutations into the *hfs* operon resulted in a 50% increase in ethanol production. We hypothesized that overexpression of this mutant *T. saccharolyticum hfsABCD* hydrogenase in  $\Delta hydG\Delta ech$  strain can be restore the redox balance by producing some hydrogen and facilitates the growth. For this purpose we constructed three expression vector using the pDGO143 backbone with three different promoter region, by Gibson assembly. We selected two promoter from *C. thermocellum* genome, *eno* and 2638, which both of the promoters are strong promoters and have been using in various gene expressions in *C. thermocellum* and we also used the native promoter of the *hfs* operon to compare the expressions. Promoter regions and the *hfs\*ABCD* gene were amplified by PCR (Figure 53) and pDGO143 digested with *Sma*I (Figure 54).

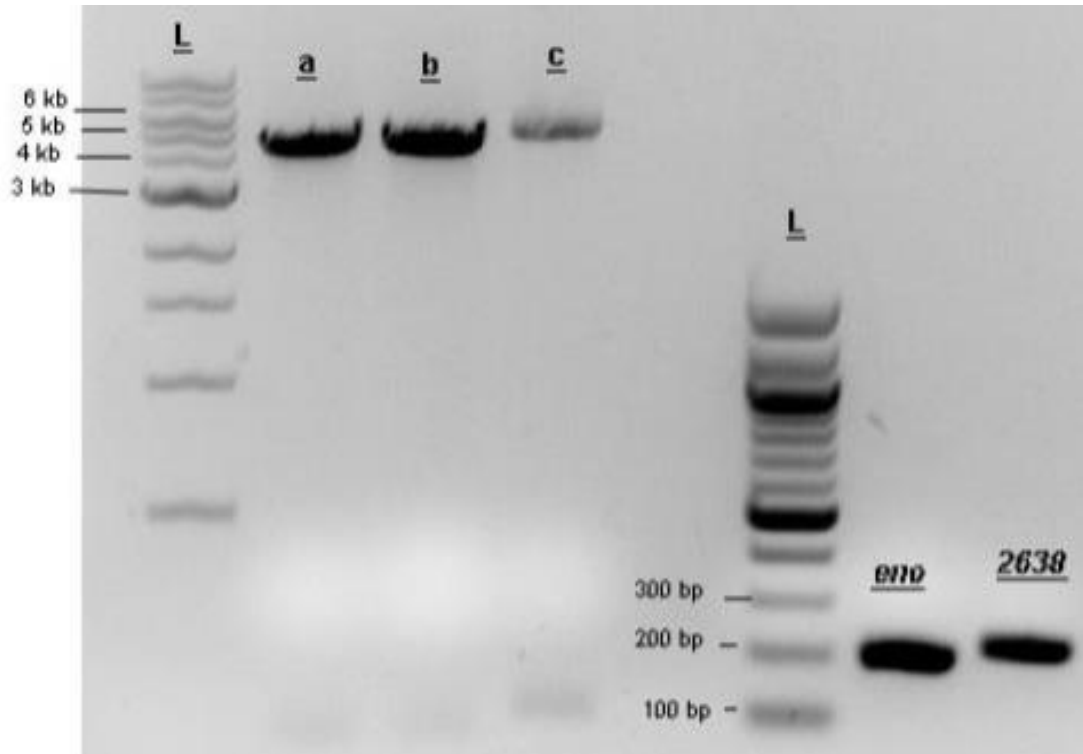


Figure 53. Fragments of *hfsABCD* and promoter regions; L: DNA ladder; a: *hfsABCD* amplified by *eno* overlapping primers; b: *hfsABCD* amplified by 2638 overlapping primers; c: *hfsABCD* with native promoter region, *eno* and 2638 are promoter fragments

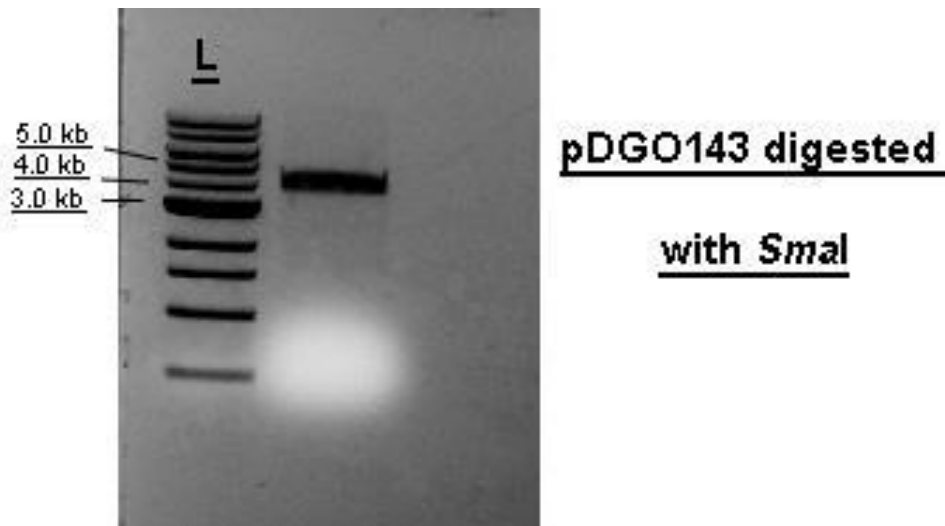


Figure 54. Linearized pDGO143 with *SmaI*

Assembled products denoted here as pDGO143- $P_{hfs}$ -*hfsABCD*\*, pDGO143- $P_{eno}$ -*hfsABCD*\* and pDGO143- $P_{2638}$ -*hfsABCD*\*, were transformed in *E. coli* NEB 5- $\alpha$  cells and positive clones were verified by colony PCR for each vector construct with an external



and internal primer pair (Figure 55, Figure 56 and Figure 57) and ~ 2900 bp region expected were obtained. Sequence verified expression vectors were transformed into the *C. thermocellum*  $\Delta hydG\Delta ech$  (LL1147). For control purposes, an empty vector was also transformed.

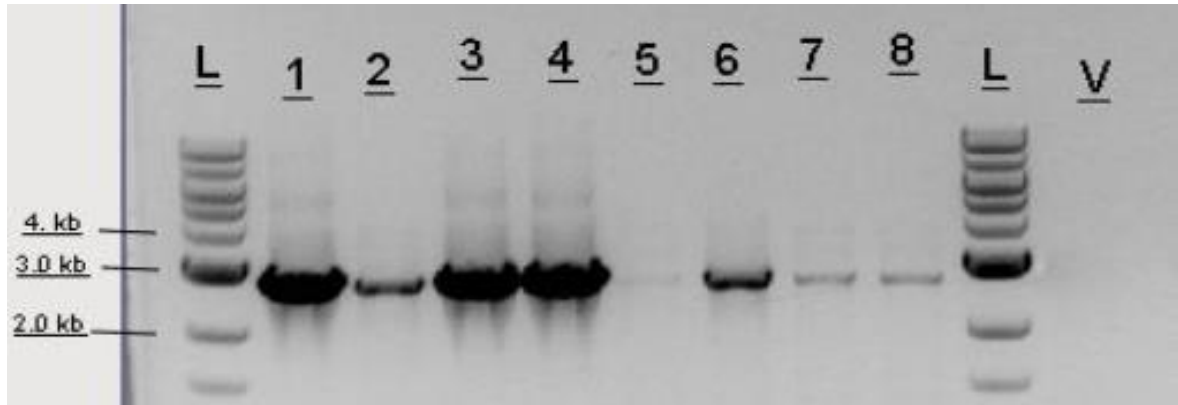


Figure 55. pDGO143- $P_{eno}$ -*hfsABCD* colony PCR with an external and internal primer; L: DNA ladder; V: vector control

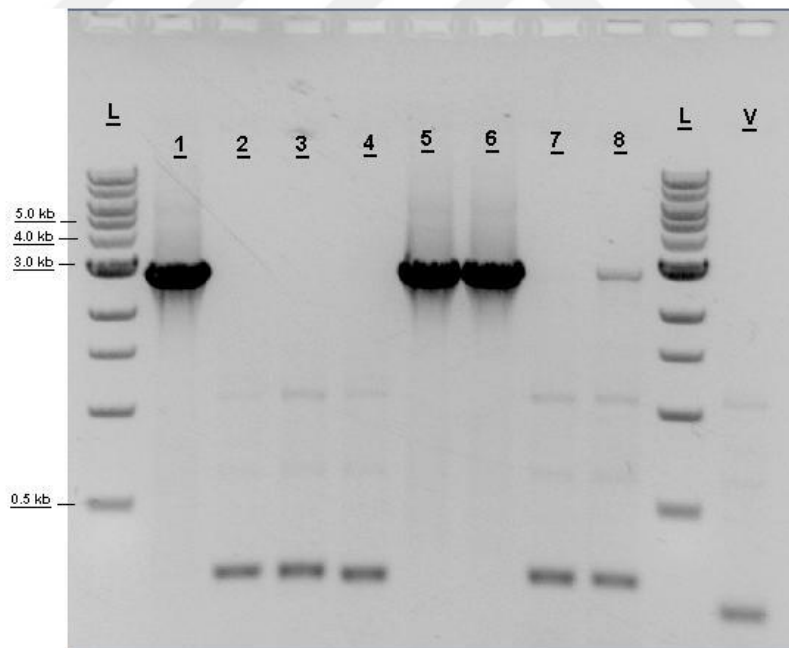


Figure 56. pDGO143- $P_{hfs}$ -*hfsABCD*\* colony PCR with an external and internal primer; L: DNA ladder; V: vector control

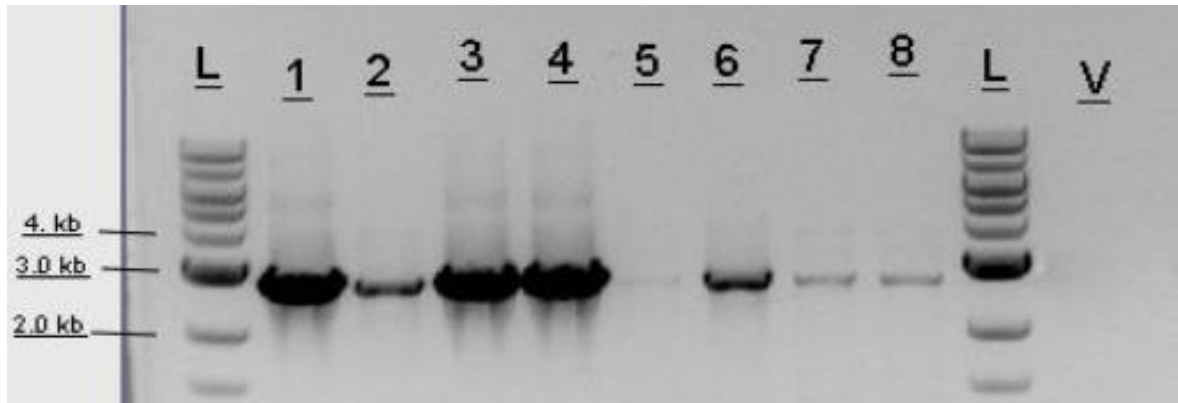


Figure 57. pDGO143-P<sub>2638</sub>-*hfsABCD*\* colony PCR with an external and internal primer; L: DNA ladder; V: vector control

After colony formation, positive strains which containing the expression vectors, were verified by colony PCR (Figure 58).

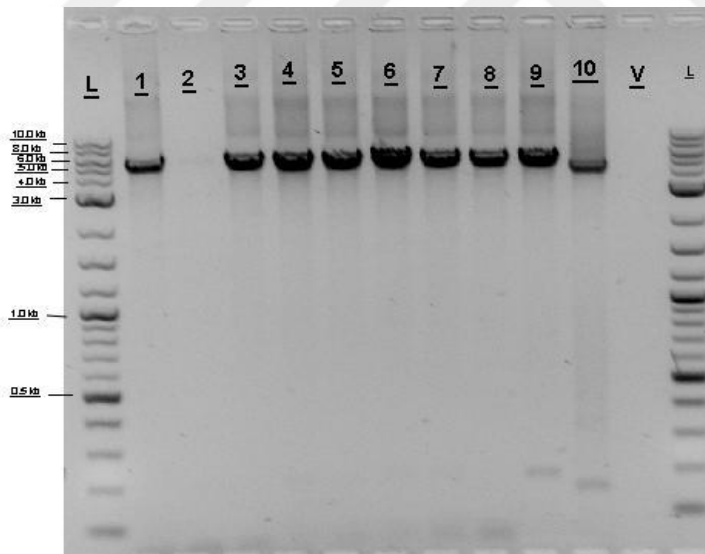


Figure 58. Colony PCR results of the LL1147 clones with an external primer pair, which are containing the expression vectors, L: DNA ladder; lines 1, 2, 3: pDGO143-P<sub>*hfs*</sub>-*hfsABCD*; lines 4, 5, 6: pDGO143-P<sub>2638</sub>-*hfsABCD* and lines 7, 8, 9: pDGO143-P<sub>*eno*</sub>-*hfsABCD*; V: vector control

### 3.13. Fermentation Profile of Overexpression of *hfsABCD* Operon in *C. thermocellum* $\Delta$ *hydGAech* (LL1147)

Overexpression of the *hfs* operon under *eno*, 2638 and native *hfs* promoters was not showed any better ethanol production. Also H<sub>2</sub> production of the overexpression mutants was lower than the parent strain LL1147 (Figure 59).

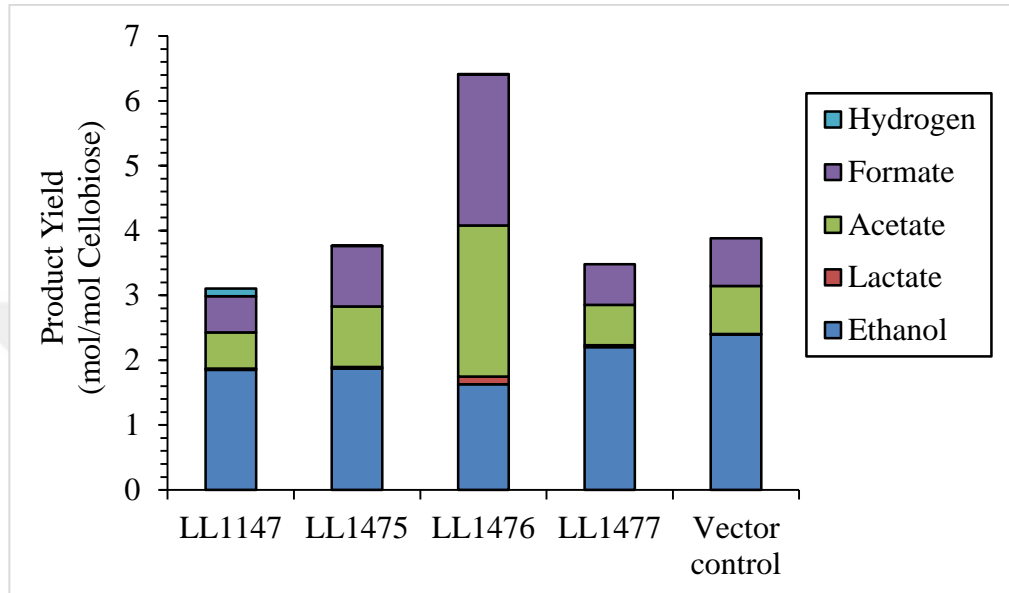


Figure 59. Fermentation products of several mutant *C. thermocellum* strains, the parent strain  $\Delta hydG\Delta ech$  (LL1147) and expression strains were grown on MTC-5 with 5 g/L cellobiose, product yield in units of moles product, produced per mole of cellobiose consumed

### 3.14. Deletion of *hfsB* in $\Delta hpt$ *C. thermocellum* DSM 1314

Removal of *hfsB* homologous gene in *C. thermocellum* gives ethanol at a molar yield of about > 2 when compared to parent strain LL345 (Figure 60) It was suggested that *hfsB* may also has a regulatory function in *C. thermocellum* on ethanol production.

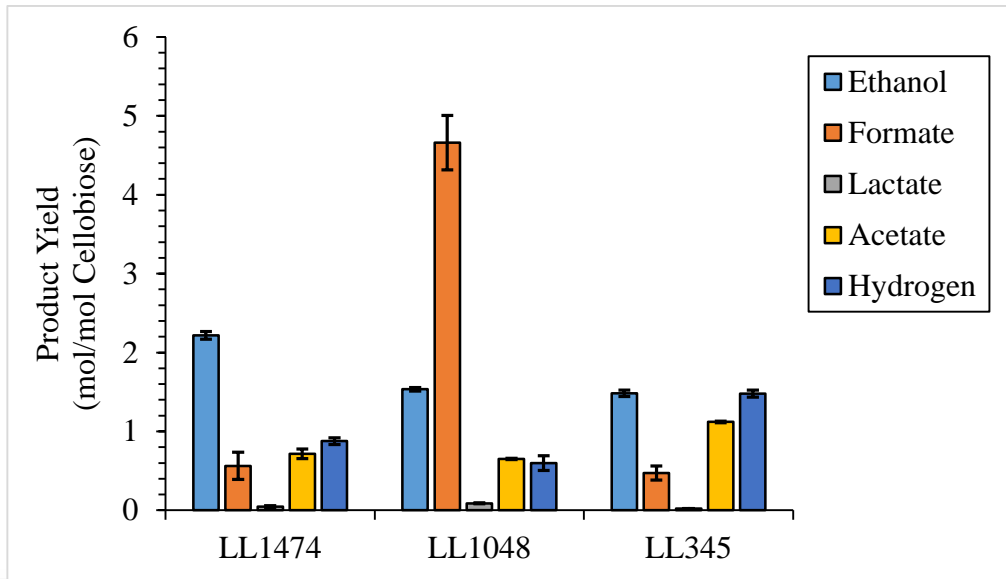


Figure 60. Fermentation products of several mutant *C. thermocellum* strains that grown on MTC-5 with 5 g/L cellobiose, product yield in units of moles product, produced per mole of cellobiose consumed, (SD=1; n=3)

#### 4. DISCUSSION

The *hfs* hydrogenase operon is a Fe-Fe hydrogenase and consisting of four subunits, *hfsA*, *hfsB*, *hfsC* and *hfsD*. Conserved domain analysis of the subunits reveals that the *hfsB* subunit contains a putative PAS domain, which function in microorganisms as redox, energy, light, oxygen, hydrogen sensors and important signalling machines in signal transduction across the proteins [117]. In this study it was aimed to understand the mechanism of the *hfs* hydrogenase operon in *T. saccharolyticum* and to determine the control mechanism of the fermentation products distribution. Identifying the function of the *hfs* hydrogenase in *T. saccharolyticum* can be helped to develop a new strategy for use in *C. thermocellum*. Both are use the same metabolic model for ethanol production. *T. saccharolyticum* is easy to manipulate on genetic level and can be useful to help understand how electron flow related to hydrogenase metabolism. Depending on previous studies it was also tried to figure out the impact of mutations in the *hfsABCD* operon. Finally it was searched an answer for the question to whether if the manipulation of the *hfs* operon can be used as a general technique with the aim of increase ethanol production in other organisms.

Other fermentation products have also been evaluated to determine the regulatory steps in this metabolic pathway. HPLC analyzes of fermentation products in mutant strains was shown that deletion of *hfsA* and/or *hfsB* subunits diverted the metabolic flux towards ethanol production, whereas deletion of the *hfsC* and/or *hfsD* subunits diverted the metabolic flux towards lactate production. Lactate is generally regarded as an overflow metabolite that is produced when the cell is under extreme stress and. Lactate dehydrogenase (LDH) is allosterically regulated by fructose 1,6 bisphosphate (FBP) in *T. saccharolyticum* [157]. LDH is inactive when cells grow rapidly because FBP levels are low but when FBP accumulates, LDH actively begins to produce lactate. Pattern of fermentation product changes is consistent with a model where *hfsA* and *B* subunits negatively regulate *hfsC* and *D* which activate ethanol production (Figure 61).

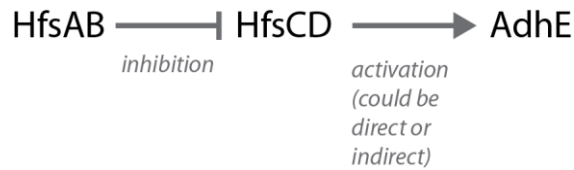


Figure 61. Hypothetical mode of action of *hfs* operon on ethanol production

The production of ethanol in *T. saccharolyticum* is carried out by AdhE and AdhA [161], so that either or both of these enzymes are thought to be the ultimate target of the regulatory pathway. When the transcriptional levels were considered, there is no relationship can be found between the deletion of the *hfs* subunits and the expression of the AdhA enzyme. Furthermore, changes in the activity of AdhA enzyme that responsible for NADPH-dependent ADH activity, found to be not related to deletion of the *hfs* subunits. Based on these evidence it was suggested that *adhA* was not participating in regulation of ethanol production by *hfs*. There was some indication that regulation of *adhE* was occurred transcriptionally. It was observed that, in *hfsA* or *B* deleted strains *adhE* expression was 2-3 fold increased, whereas in *hfsC* or *D* deleted strains *adhE* expression was not changed. Even so the difference in the *adhE* expression was not sufficient to explain the regulation mechanism. As a bifunctional enzyme AdhE has two activity; ADH and ALDH activities [20] and in wt *T. saccharolyticum*, these activities are dependent on the NADH. If the increase in ethanol production was due to the observed increase in the expression of *adhE*, NADH-dependent ADH and ALDH activities would be expected to increase in similar amounts. However, it was observed that, NADH-ADH activity was increased (p=0.005), but NADH-ALDH activity was not changed. For these reasons, a possible regulatory mechanism between HFSAB, CD and ADHE emerges as the most likely model in this work.

There also some other possibilities which might be considered for the regulation through signal transduction between *hfsCD* and *adhE*. In many organisms it was shown that in response to changes in the NADH/NAD<sup>+</sup> ratio the *rex* gene which is a redox sensing transcriptional regulator, regulate *adhE* expression [116]. The Rex protein was bound to DNA in the upstream region of a gene, inhibiting its expression and thus preventing transcription but if an excess of NADH was present, it break downs the Rex-DNA binding following with transcription of the target gene.

In the light of these data, it is expected that the deletion of the *rex* gene increases the production of ethanol and vice versa but the expression analysis obtained in this study

showed that the expression of the *rex* gene also increased in strains with high ethanol production.

Another gene that may be effective in the regulatory pathway is *tsac\_0415*. This gene shows the highest variability in the expression among the high and low ethanol producing strains. This gene is also adjacent to the *adhE* gene on the genome. A BLAST search revealed only 5 close matches in the nr database, all in the *Thermoanaerobacterium* genus including *T. xyloxyticum* and *T. thermosaccharolyticum* [8] even it is annotated by PFAM as “unknown function” [162]. To confirm this hypothesis further experimental evidence was needed.

The elimination of lactic acid as a fermentation product resulted in an increase in acetic acid and ethanol yield, as previously reported by Saw et al. [16]. Regarding to this data, it has been suggested that, deletion of the *ldh* in the  $\Delta hfsA::kan$  and  $\Delta hfsB::kan$  strains could have resulted in an increase in ethanol production via complete elimination of the *ldh* gene. To test this hypothesis the *ldh* gene was replaced with a ~3.4 kb *erm* resistance cassette in LL1267 and LL1268. After the fermentation data analysis, we couldn't see any better increase in ethanol production even there were almost no lactate production in those two strains. This results indicated that the deletion of the *hfsA* and/or *hfsB* also lead to downregulation of *ldh* in *T. saccharolyticum*.

*C. thermocellum* is a thermophilic, cellulolytic bacterium that can natively solubilize cellulose and produce ethanol as a fermentation product however, in wild-type *C. thermocellum* ethanol yield and titer is low and additionally it is producing acetate, lactate, H<sub>2</sub>, formate, and free amino acids. Different genetic manipulation approaches have been developed to obtain high ethanol yield in this bacteria.

In a previous study hydrogenase active site assembly was eliminated and this blocked H<sub>2</sub> production, by simultaneous deletion of the; hydrogenase maturase gene (*hydG*) and the [NiFe] hydrogenase (*ech*), increased ethanol yield in *C. thermocellum* nearly to 64% of the theoretical maximum. However, the new strain was grown slowly, due to difficulty balancing redox potential and was less ethanol tolerant [128].

In a different study it has been reported that the *hfs* hydrogenase is responsible for H<sub>2</sub> production in *T. saccharolyticum*. Deletion of the *hfs* operon resulted in a decrease in H<sub>2</sub> and acetate, but no change in ethanol production. Introduction of two point mutations into the *hfs* operon resulted in a 50% increase in ethanol formation We have hypothesized that overexpression of the mutant *T. saccharolyticum*  $\Delta hfsABCD::kan$  hydrogenase (*hfs\**) in

*ΔhydGΔech* strain of the *C. thermocellum* can restore the redox balance by producing some H<sub>2</sub> and facilitates the growth. However the overexpression of the *hfs* operon in the *ΔhydGΔech C. thermocellum* strain did not result with an increase in ethanol yield.

Consequently the overexpression of the *hfs* operon in the *ΔhydGΔech C. thermocellum* strain was not resulted with an increase in ethanol yield. Deleting the *hfsB* gene (which predicted as a histidine kinase containing a PAS domain that probable functioning as redox or hydrogen sensing apparatus and a hydrogenase domain) in wild type *C. thermocellum* resulted at a molar yield of about >2 in ethanol production.





## 5. CONCLUSION

Previously all 4 subunits had been deleted simultaneously. In this work, each subunit disrupted individually by replacing with a kanamycin antibiotic resistance marker. After HPLC analysis of the fermentation products it was seen that when all 4 subunits of *hfs* operon are deleted (A, B, C and D, strain LL1187), ethanol production is lower than wild-type. When either *hfsA* or *hfsB* is deleted, but *hfsC* and *hfsD* are intact, ethanol production is similar to ethanologen strains.

To better understand the mechanism for increased ethanol production in  $\Delta hfsB$  strain of *T. saccharolyticum* qPCR, RNAseq, and proteomics analyses were performed. As a result of the OMICS analysis, significant (4-5x) upregulation of *adhE* at both gene and protein level was observed. The whole genome sequencing results corroborate to the idea on the deletion of the *hfsA* or *hfsB* might be a useful approach for metabolic engineering some thermophilic bacteria for increased ethanol production.

Although it was shown that in *T. saccharolyticum* deletions of *hfsA* and/or *B* are sufficient to obtain a high-yielding ethanologenic strain, it has been investigated whether this manipulation technique is also useful in other organisms. It was found that deletion of *hfsB* in *T. xylandolyticum* and *T. thermosaccharolyticum* cause in large increases in ethanol production, but in other strains the increase in ethanol production was proportionally less than the *Thermoanaerobacterium* strains. As a result, it was shown that the deletion of *hfsB* is can be used as a genetic engineering technique to increase ethanol production in a variety of organisms.

In this work a metabolic mechanism that function to regulate the distribution of fermentation products was presented. Evidence form proteomics analysis and expression analysis data support the regulator role of *hfsB* on *adhE* which is essential in ethanol production in these organisms. This new fundings has shown that *hfs* operon is function as a regulatory pathway in these organism on ethanol production and showed that deleting only one gene (*hfsB*) is sufficient for developing ethanologen strains in *T. saccharolyticum* and in several other microorganisms.

## 6. RECOMMENDATIONS

Since the beginning of this research, studies have shown that ethanol production with a single gene deletion in thermophilic bacteria containing the *hfs* operon can be increased at the theoretical maximum of 78%. It could be important that the subsequent work will be carried out on the full characterization of this regulation mechanism. In this frame, the functions of the expression products of the genes that constitute *T. saccharolyticum* Hfs operon must first be elucidated.

To determine whether the PAS domain in *hfsB* is involved in the predicted regulation mechanism, it may be necessary to examine only the deletion of this domain and its effect on ethanol production in the gland.

In BLAST analyzes performed with the amino acid sequences of *hfsB* and *hfsC* in NCBI, it was predicted that the *hfsB* protein could be a possible histidine kinase and the *hfsC* protein could be a possible serine / threonine phosphatase. Understanding whether the regulation is via protein phosphorylation may play an important role in clearly revealing this regulation mechanism.

## 7. REFERENCES

1. Wyman, C., E., Biomass ethanol: Technical Progress, Opportunities, and Commercial Challenges, Annual Review of Energy and the Environment, 24,1 (1999) 189-226.
2. <https://ec.europa.eu/energy/en/topics/renewable-energy>. Renewable energy. 31.10.2017.
3. Scully, S., M. and Orlygsson, J., Recent Advances in Second Generation Ethanol Production by Thermophilic Bacteria, Energies, 8,1 (2014) 1-30.
4. Van Der Veen, D., Lo, J., Brown, S., D., Johnson, C., M., Tschaplinski, T., J., Martin, M., Engle, N., L., Van den Berg, R., A., Argyros, A., D., Caiazza, N., C. and Guss, A., M., Characterization of *Clostridium thermocellum* Strains with Disrupted Fermentation End-product Pathways, Journal of Industrial Microbiology and Biotechnology, 40,7 (2013) 725-734.
5. Lynd, L., R., Van Zyl, W., H., McBride, J., E. and Laser, M., Consolidated Bioprocessing of Cellulosic Biomass: An Update, Current Opinion in Biotechnology, 16,5 (2005) 577-583.
6. NREL/BR, National Renewable Energy Laboratory Research Review (2003-2004) Publication no: 840-36178, Golden, CO, 2003.
7. Olson, D., G., McBride, J., E., Shaw, A., J. and Lynd, L., R., Recent Progress in Consolidated Bioprocessing, Current Opinion in Biotechnology, 23,3 (2012) 396-405.
8. Lynd, L., R., Weimer, P., J., Van Zyl, W., H. and Pretorius, I., S., Microbial Cellulose Utilization: Fundamentals and Biotechnology, Microbiology and Molecular Biology Reviews, 66,3 (2002) 506-577.
9. Chang, T. and Yao, S., Thermophilic, Lignocellulolytic Bacteria for Ethanol Production: Current State and Perspectives, Applied Microbiology and Biotechnology, 92,1 (2011) 13-27
10. Olson, D., G. and Lynd, L., R., Transformation of *Clostridium thermocellum* by Electroporation, Methods in Enzymology, 510 (2012) 317-330.
11. Olson, D., G., Maloney, M., Lanahan, A., A., Hon, S., Hauser, L., J. and Lynd, L., R., Identifying Promoters for Gene Expression in *Clostridium thermocellum*. Metabolic Engineering Communication, 2 (2015) 23-29.
12. Sims, R., E., Mabee, W., Saddler, J., N. and Taylor, M., An Overview of Second Generation Biofuel Technologies, Bioresource Technology, 101,6 (2010), 1570-1580.

13. Singer, S., Denruyter J., P., Jeffries, B., The Energy Report 100% Renewable Energy by 2025. <https://www.ecofys.com/files/files/ecofys-wwf-2011-the-energy-report.pdf> December 10, 2017
14. Dale, B., E., Anderson, J., E., Brown, R., C., Csonka, S., Dale, V., H., Herwick, G. and Wang, M., Q., Take a Closer Look: Biofuels can Support Environmental, Economic and Social Goals, Environmental Science and Technology, 48,13 (2014) 7200-7203.
15. Margeot, A., Hahn-Hagerdal, B., Edlund, M., Slade, R. and Monot, F., New Improvements for Lignocellulosic Ethanol, Current Opinion in Biotechnology, 20,3 (2009) 372-380.
16. Olson, D., G., Sparling, R. and Lynd, L., R., Ethanol Production by Engineered Thermophiles, Current Opinion in Biotechnology, 33 (2015) 130-141.
17. Shaw, A., J., Podkaminer, K., K., Desai, S., G., Bardsley, J., S., Rogers, S., R., Thorne, P., G. and Lynd, L., R., Metabolic Engineering of a Thermophilic Bacterium to Produce Ethanol at High Yield, Proceedings of the National Academy of Sciences, 105,37 (2008) 13769-13774.
18. Zhou, J., Olson, D., G., Argyros, D., A., Deng, Y., van Gulik, W., M., van Dijken, J., P. and Lynd, L., R., Atypical Glycolysis in *Clostridium thermocellum*, Applied and Environmental Microbiology, 79,9 (2013) 3000-3008.
19. Tyurin, M., V., Desai, S., G. and Lynd, L., R., Electrotransformation of *Clostridium thermocellum*, Applied and Environmental Microbiology, 70,2 (2004) 883-890.
20. Tripathi, S., A., Olson, D., G., Argyros, D., A., Miller, B., B., Barrett, T., F., Murphy, D., M. and Caiazza, N., C., Development of pyrF-based Genetic System for Targeted Gene Deletion in *Clostridium thermocellum* and Creation of A *pta* Mutant, Applied and Environmental Microbiology, 76,19 (2010) 6591-6599.
21. Argyros, D., A., Tripathi, S., A., Barrett, T., F., Rogers, S., R., Feinberg, L., F., Olson, D., G. and Caiazza, N., C., High Ethanol Titters from Cellulose by Using Metabolically Engineered Thermophilic, Anaerobic Microbes, Applied and Environmental Microbiology, 77,23 (2011) 8288-8294.
22. Lin, P., P., Mi, L., Morioka, A., H., Yoshino, K., M., Konishi, S., Xu, S., C. and Liao, J., C., Consolidated Bioprocessing of Cellulose to Isobutanol Using *Clostridium thermocellum*, Metabolic Engineering, 31 (2015) 44-52.
23. Barnard, D., Casanueva, A., Tuffin, M. and Cowan, D., Extremophiles in Biofuel Synthesis, Environmental Technology, 31,8-9 (2010) 871-888.
24. Shaw, A., J., Hogsett, D., A. and Lynd, L., R., Natural Competence in *Thermoanaerobacter* and *Thermoanaerobacterium* Species, Applied and Environmental Microbiology, 76,14 (2010) 4713-4719.

25. Tyurin, M., V., Lynd, L., R. and Wiegel, J., 13 Gene Transfer Systems for Obligately Anaerobic Thermophilic Bacteria, Methods in Microbiology, 35 (2006) 309-330.
26. Ganghofner, D., Kellermann, J., Staudenbauer, W., L. and Bronnenmeier, K., Purification and Properties of An Amylopullulanase, A Glucoamylase, and An  $\alpha$ -Glucosidase in the Amyolytic Enzyme System of *Thermoanaerobacterium thermosaccharolyticum*, Bioscience Biotechnology and Biochemistry, 62,2 (1998) 302-308.
27. Lee, Y., E., Jain, M., K., Lee, C. and Zeikus, J., G., Taxonomic Distinction of Saccharolytic Thermophilic Anaerobes: Description of *Thermoanaerobacterium xylanolyticum* gen. nov., sp. nov., and *Thermoanaerobacterium saccharolyticum* gen. nov., sp. nov.; Reclassification of *Thermoanaerobium brockii*, *Clostridium thermosulfurogenes*, and *Clostridium thermohydrosulfuricum* E100-69 as *Thermoanaerobacter brockii* comb. nov., *Thermoanaerobacterium thermosulfurigenes* comb. nov., and *Thermoanaerobacter thermohydrosulfuricus* comb. nov., Respectively; and Transfer of *Clostridium thermohydrosulfuricum* 39E to *Thermoanaerobacter ethanolicus*, International Journal of Systematic and Evolutionary Microbiology, 43,1 (1993) 41-51.
28. Shaw, A., J., Jenney, F., E., Adams, M., W. and Lynd, L., R., End-product Pathways in the Xylose Fermenting Bacterium, *Thermoanaerobacterium saccharolyticum*, Enzyme and Microbial Technology, 42,6 (2008) 453-458.
29. Desai, S., G., Guerinot, M., L. and Lynd, L., R., Cloning of L-lactate Dehydrogenase and Elimination of Lactic Acid Production via Gene Knockout in *Thermoanaerobacterium saccharolyticum* JW/SL-YS485, Applied and Environmental Microbiology, 65,5 (2004) 600-605.
30. Shaw, A., J., Miller, B., B., Rogers, S., R., Kenealy, W., R., Meola, A., Bhandiwad, A. and Herring, C., D., Anaerobic Detoxification of Acetic Acid in A Thermophilic Ethanologen, Biotechnology for Biofuels, 8,1 (2015) 75.
31. Shaw, A., J., Hogsett, D., A. and Lynd, L., R., Characterization of the [FeFe]-hydrogenase Responsible for Hydrogen Generation in *Thermoanaerobacterium saccharolyticum* and Demonstration of Increased Ethanol Yield via Hydrogenase Knockout, Journal of Bacteriology, 191,20 (2009) 6457-6464.
32. Herring, C., D., Kenealy, W., R., Shaw, A., J., Covalla, S., F., Olson, D., G., Zhang, J. and Thorne, P., G., Strain and Bioprocess Improvement of A Thermophilic Anaerobe for the Production of Ethanol from Wood, Biotechnology for Biofuels, 9,1 (2016) 125.
33. Rosillo-Calle, F. and Woods, J., The Biomass Assessment Handbook: Bioenergy for A Sustainable Environment, First Edition, Earthscan, London, 2012.
34. McKendry, P., Energy Production from Biomass (part 1): Overview of Biomass, Bioresource Technology, 83,1 (2002) 37-46.

35. Duku, M., H., Gu, S. and Hagan, E., B., A Comprehensive Review of Biomass Resources and Biofuels Potential in Ghana, Renewable and Sustainable Energy Reviews, 15,1 (2011) 404-415.
36. Scarlat, N., Dallemand, J., F., Monforti-Ferrario, F. and Nita, V., The Role of Biomass and Bioenergy in A Future Bioeconomy: Policies and Facts, Environmental Development, 15 (2015) 3-34.
37. World Bioenergy Association (WBA), Global Biomass Potential Towards 2035, Stockholm, March 2016.
38. Breeze, P., The Future of Global Biomass Power Generation, Business Insights Ltd., London, 2004.
39. Serrano-Ruiz, J., C., West, R., M. and Dumesic, J., A., Catalytic Conversion of Renewable Biomass Resources to Fuels and Chemicals, Annual Review of Chemical and Biomolecular Engineering, 1 (2010) 79-100.
40. Malherbe, S. and Cloete T., E., Lignocellulose Biodegradation: Fundamentals and Applications, Reviews in Environmental Science and Bio/Technology, 1,2 (2002) 105-114.
41. Howard, R., L., Abotsi, E., L. J., R., Van Rensburg, E., J. and Howard, S., Lignocellulose Biotechnology: Issues of Bioconversion and Enzyme Production, African Journal of Biotechnology, 2,12 (2003) 602-619.
42. Knauf, M. and Moniruzzaman M., Lignocellulosic Biomass Processing: A Perspective, International Sugar Journal, 106,1263 (2004) 147-150.
43. Menon, V. and Rao M., Trends in Bioconversion of Lignocellulose: Biofuels, Platform Chemicals & Biorefinery Concept, Progress in Energy and Combustion Science, 38,4 (2012) 522-550.
44. Isikgor, F., H. and Becer, C., R., Lignocellulosic Biomass: A Sustainable Platform for the Production of Bio-based Chemicals and Polymers, Polymer Chemistry, 6,25 (2015) 4497-4559.
45. Saha, B.C., Enzymes as Biocatalysts for Conversion of Lignocellulosic Biomass to Fermentable Sugars, Handbook of Industrial Biocatalysis, (2005) 1-12.
46. Pratima, B., Pretreatment of lignocellulosic biomass for biofuel production, First Edition Springer, Singapore, 2016.
47. Guerriero, G., Hausman, J., F., Strauss, J., Ertan, H. and Siddiqui, K., S., Lignocellulosic Biomass: Biosynthesis, Degradation, and Industrial Utilization, Engineering in Life Sciences, 16,1 (2016) 1-16.

48. Akil, H., Omar, M., F., Mazuki, A., A., M., Safiee, S., Z., A., M., Ishak, Z., M. and Bakar, A., A., Kenaf Fiber Reinforced Composites: A Review, Materials & Design, 32,(8-9) (2011) 4107-4121.
49. Laureano-Perez L., Teymouri F., Alizadeh H. and Dale B., E., Understanding Factors That Limit Enzymatic Hydrolysis of Biomass: Characterization of Pretreated Corn Stover, Applied Biochemistry and Biotechnology, 124,(1-3) (2005) 1081–1099.
50. Fengel, D. and Wegener, G., Wood Chemistry, Ultrastructure, Reactions, Walter de Gruyter, Berlin, 1984.
51. Lee, H., V., Hamid, S., B. A. and Zain, S., K., Conversion of Lignocellulosic Biomass to Nanocellulose: Structure and Chemical Process, The Scientific World Journal 2014 (2014).
52. Howard, R., L., Abotsi, E., L., J, R., Van Rensburg, E., J. and Howard, S., Lignocellulose biotechnology: Issues of Bioconversion and Enzyme Production, African Journal of Biotechnology, 2,12 (2003) 602-619.
53. Lee, H., V., Hamid, S., B., A. and Zain, S., K.. Conversion of Lignocellulosic Biomass to Nanocellulose: Structure and Chemical Process, The Scientific World Journal, 2014 (2014).
54. Kumar, R., Singh, S. and Singh, O., V., Bioconversion of Lignocellulosic Biomass: Biochemical and Molecular Perspectives, Journal of Industrial Microbiology & Biotechnology, 35,5 (2008) 377-391.
55. Ben-Iwo, J., Manovic, V. and Longhurst, P., Biomass resources and biofuels potential for the production of transportation fuels in Nigeria, Renewable and Sustainable Energy Reviews, 63 (2016) 172-192.
56. Lange, J., P., Lignocellulose Conversion: An Introduction to Chemistry, Process and Economics, Biofuels, Bioproducts and Biorefining, 1,1 (2007) 39-48.
57. Iglesia, E., Reyes, S., C., Madon, R., J. and Soled, S., L., Selectivity Control and Catalyst Design in the Fischer-Tropsch Synthesis: Sites, Pellets, and Reactors, Advances In Catalysis, 39 (1993) 221-302.
58. Caldwell, L., Selectivity in Fischer-Tropsch Synthesis: Review and Recommendations for Further work, 1980.
59. Klier, K., Methanol synthesis, Advances in Catalysis, 31 (1982) 243-313.
60. Lange, J., P., Methanol Synthesis: A Short Review of Technology Improvements." Catalysis Today, 64,1 (2001) 3-8.
61. Kunkes, E., L., Soares, R., R., Simonetti, D., A. and Dumesic, J., A., An Integrated Catalytic Approach For the Production of Hydrogen by Glycerol Reforming Coupled with Water-Gas Shift, Applied Catalysis B: Environmental, 90,(3-4) (2009) 693-698.

62. Soares, R., R., Simonetti, D., A. and Dumesic, J., A., Glycerol as A Source for Fuels and Chemicals by Low-Temperature Catalytic Processing, Angewandte Chemie, 118,24 (2006) 4086-4089.
63. Huber, G., W., Iborra, S. and Corma, A., Synthesis of Transportation Fuels from Biomass: Chemistry, Catalysts, and Engineering, Chemical Reviews, 106,9 (2006) 4044-4098.
64. Spath, P., L. and Dayton, D., C., Preliminary Screening-Technical and Economic Assessment of Synthesis Gas to Fuels and Chemicals with Emphasis On the Potential for Biomass-Derived Syngas.. National Renewable Energy Lab, Publication No: NREL/TP-510-34929, Golden Co, 2003.
65. Lange, J., P., Lignocellulose Conversion: An Introduction to Chemistry, Process and Economics, Biofuels, Bioproducts and Biorefining, 1,1 (2007) 39-48.
66. Mohan, D., Pittman, C., U. and Steele, P., H., Pyrolysis of wood/biomass for bio-oil: a critical review, Energy & Fuels 20,3 (2006) 848-889.
67. Breaking the Chemical and Engineering Barriers to Lignocellulosic Biofuels and Huber, G., W., Breaking the Chemical and Engineering Barriers to Lignocellulosic Biofuels: Next Generation Hydrocarbon Biorefineries, National Science Foundation, Chemical, Biogengineering, Environmental and Transport Systems Division, Washington, DC, 2008.
68. Mohan, D., Pittman, C., U. and Steele, P., H., Pyrolysis of Wood/Biomass for Bio-Oil: A Critical Review, Energy & Fuels, 20,3 (2006) 848-889.
69. Elliott, D., C., Historical Developments In Hydroprocessing Bio-Oils, Energy & Fuels, 21,3 (2007) 1792-1815.
70. Kersten, S., R., A., van Swaaij, W., P., M., Lefferts, L. and Seshan, K., Options for Catalysis in the Thermochemical Conversion of Biomass into Fuels, in Catalysis for Renewables: From Feedstock to Energy Production, Wiley-VCH Verlag GmbH & Co., KGaA, Weinheim, Germany, 2007.
71. Carlson, T., R., Vispute T., P. and Huber, G., W., Green Gasoline By Catalytic Fast Pyrolysis of Solid Biomass Derived Compounds, Chem Sus Chem, 1,5 (2008) 397-400.
72. Nigam, P., S. and Singh, A., Production of liquid biofuels from renewable resources, Progress In Energy and Combustion Science, 37,1 (2011) 52-68.
73. Anwar, Z., Gulfranz, M. and Irshad, M., Agro-Industrial Lignocellulosic Biomass A Key to Unlock the Future Bio-Energy: A Brief Review, Journal of Radiation Research and Applied Sciences, 7,2 (2014) 163-173.



74. Wyman, C. E., Dale, B., E., Elander, R., T., Holtzapple, M., Ladisch, M., R. and Lee, Y., Y., Coordinated Development of Leading Biomass Pretreatment Technologies, Bioresource Technology, 96,18 (2005) 1959-1966.
75. Saini, J., K., Saini, R. and Tewari, L., Lignocellulosic Agriculture Wastes as Biomass Feedstocks for Second-Generation Bioethanol Production: Concepts and Recent Developments, 3 Biotech, 5,4 (2015) 337-353.
76. Lynd, L., R., Overview and Evaluation of Fuel Ethanol from Cellulosic Biomass: Technology, Economics, the Environment, and Policy, Annual Review of Energy and the Environment, 21,1 (1996) 403-465.
77. Sun, Y. and Cheng, J., Hydrolysis of Lignocellulosic Materials for Ethanol Production: A Review, Bioresource Technology, 83,1 (2002) 1-11.
78. Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y., Y., Holtzapple, M. and Ladisch, M., Features of Promising Technologies for Pretreatment of Lignocellulosic Biomass, Bioresource Technology, 96,6 (2005) 673-686.
79. Duff, S., J. and Murray, W., D., Bioconversion of Forest Products Industry Waste Cellulosics to Fuel Ethanol: A Review. Bioresource Technology, 55,1 (1996) 1-33.
80. Bisaria, V., S., Bioprocessing of Agro-Residues to Glucose and Chemicals, Bioconversion of Waste Materials to Industrial Products, (1991) 187-223.
81. Singh, A., Kumar, P., K., R. and Schügerl, K., Bioconversion of Cellulosic Materials to Ethanol by Filamentous Fungi, Enzymes and Products from Bacteria Fungi and Plant Cells, Springer, Berlin, 1992.
82. Prasad, S., Singh, A. and Joshi, H., C., Ethanol as An Alternative Fuel from Agricultural, Industrial and Urban Residues, Resources, Conservation and Recycling, 50,1 (2007) 1-39.
83. Álvarez, C., Reyes-Sosa, F., M. and Díez, B., Enzymatic Hydrolysis of Biomass from Wood, Microbial Biotechnology, 9,2 (2016) 149-156.
84. Lynd, L., R., Currie, D., Ciazza, N., Herring, C., and Orem, N., Consolidated Bioprocessing of Cellulosic Biomass to Ethanol Using Thermophilic Bacteria, American Society of Microbiology, January 2008, In Bioenergy, 55-74
85. Lynd, L., R., Overview and Evaluation of Fuel Ethanol From Cellulosic Biomass: Technology, Economics, the Environment, and Policy, Annual Review of Energy and the Environment, 21,1 (1996) 403-465.
86. Avgerinos, G., C., Biocic, I., Wang, S., D. and Fang, H., Y., Ethanol from cellulosic biomass, Philosophical Transactions of the Royal Society of London Series B, 300,1100 (1983) 323-333.

87. Esterbauer, H., Steiner, W., Labudova, I., Hermann, A. and Hayn, M., Production of Trichoderma Cellulase in Laboratory and Pilot Scale, Bioresource Technology, 36,1 (1991) 51-65.
88. Wright, J., D., Wyman, C., E. and Grohmann, K., Simultaneous Saccharification and Fermentation of Lignocellulose, Applied Biochemistry and Biotechnology, 18,1 (1988) 75-90.
89. Wooley, R., Ruth, M., Glassner, D. and Sheehan, J., Process Design and Costing of Bioethanol Technology: A Tool for Determining the Status and Direction of Research and Development, Biotechnology Progress, 15,5 (1999) 794-803.
90. Lynd, L., R., Cushman, J., H., Nichols, R., J. and Wyman, C., E., Fuel Ethanol from Cellulosic Biomass, Science, 251,4999 (1991) 1318-1323.
91. Lynd, L., R., Elander, R., T. and Wyman, C., E., Likely Features and Costs of Mature Biomass Ethanol Technology, Seventeenth Symposium on Biotechnology for Fuels and Chemicals, Humana Press, Totowa, NJ, 1996.
92. Ingram, L., O., Aldrich, H., C., Borges, A., C. C., Causey, T., B., Martinez, A., Morales, F. and Zaldivar, J., Enteric Bacterial Catalysts For Fuel Ethanol Production, Biotechnology Progress, 15,5 (1999) 855-866.
93. Ingram, L., O., Gomez, P., F., Lai, X., Moniruzzaman, M., Wood, B., E., Yomano, L., P. and York, S., W., Metabolic Engineering of Bacteria for Ethanol Production, Biotechnology and Bioengineering, 58,2-3 (1998) 204-214.
94. Ho, N., W., Chen, Z., Brainard, A., P. and Sedlak, M., Successful Design and Development of Genetically Engineered *Saccharomyces* Yeasts for Effective Cofermentation of Glucose and Xylose from Cellulosic Biomass to Fuel Ethanol, Recent Progress in Bioconversion of Lignocellulosics, Springer, Berlin, Heidelberg, 1999
95. Hahn-Hägerdal, B., Wahlbom, C., F., Gárdonyi, M., van Zyl, W., H., Otero, R., R. C. and Jönsson, L., J., Metabolic Engineering of *Saccharomyces cerevisiae* for Xylose Utilization, Metabolic Engineering, (2001) 53-84.
96. Deanda, K., Zhang, M., Eddy, C. and Picataggio, S., Development of An Arabinose-Fermenting *Zymomonas mobilis* Strain by Metabolic Pathway Engineering, Applied and Environmental Microbiology, 62,12 (1996) 4465-4470.
97. Zhang, M., Eddy, C., Deanda, K., Finkelstein, M. and Picataggio, S., Metabolic Engineering of A Pentose Metabolism Pathway in Ethanologenic *Zymomonas mobilis*, Science, 267,5195 (1995) 240-243.
98. Taylor, M., P., Eley, K., L., Martin, S., Tuffin, M., I., Burton, S., G. and Cowan, D., A., Thermophilic Ethanogenesis: Future Prospects for Second-Generation Bioethanol Production, Trends in Biotechnology, 27,7 (2009) 398-405.

99. Lynd, L., R. and Zhang, Y., Quantitative Determination of Cellulase Concentration as Distinct from Cell Concentration in Studies of Microbial Cellulose Utilization: Analytical Framework and Methodological Approach, Biotechnology and Bioengineering, 77,4 (2002) 467-475.
100. Patel, M., A., Ou, M., S., Ingram, L., O. and Shanmugam, K., T., Simultaneous Saccharification and Co-Fermentation of Crystalline Cellulose and Sugar Cane Bagasse Hemicellulose Hydrolysate to Lactate by A Thermotolerant Acidophilic *Bacillus* sp, Biotechnology Progress, 21,5 (2005) 1453-1460.
101. Bhandiwad, A., Shaw, A., J., Guss, A., Guseva, A., Bahl, H. and Lynd, L., R., Metabolic Engineering of *Thermoanaerobacterium saccharolyticum* for n-Butanol Production, Metabolic Engineering, 21 (2014) 17-25.
102. Liu, S., Y., Gherardini, F., C., Matuschek, M., Bahl, H. and Wiegel, J., Cloning, Sequencing, and Expression of the Gene Encoding A Large S-Layer-Associated Endoxylanase from *Thermoanaerobacterium* sp. strain JW/SL-YS 485 in *Escherichia coli*, Journal of Bacteriology, 178,6 (1996) 1539-1547.
103. Mai, V. and Wiegel, J., Advances in Development of a Genetic System for *Thermoanaerobacterium* spp.: Expression of Genes Encoding Hydrolytic Enzymes, Development of a Second Shuttle Vector, and Integration of Genes into the Chromosome, Applied and Environmental Microbiology, 66,11 (2000) 4817-4821.
104. Shaw, A., J., Covalla, S., F., Hogsett, D., A. and Herring, C., D., Marker Removal System for *Thermoanaerobacterium saccharolyticum* and Development of A Markerless Ethanologen, Applied and Environmental Microbiology, 77,7 (2011) 2534-2536.
105. Lynd, L., R., Baskaran, S. and Casten, S., Salt Accumulation Resulting from Base Added for pH Control, and Not Ethanol, Limits Growth of *Thermoanaerobacterium thermosaccharolyticum* HG-8 at Elevated Feed Xylose Concentrations in Continuous Culture, Biotechnology Progress, 17,1 (2001) 118-125.
106. Shaw, A., J., Covalla, S., F., Miller, B., B., Firliet, B., T., Hogsett, D., A. and Herring, C., D., Urease Expression in a *Thermoanaerobacterium saccharolyticum* Ethanologen Allows High Titer Ethanol Production, Metabolic Engineering, 14,5 (2012) 528-532.
107. Herring, C., D., Kenealy, W., R., Shaw, A., J., Raman, B., Tschaplinski, T., J. and Brown, S. D., Final report on development of *Thermoanaerobacterium saccharolyticum* for the conversion of lignocellulose to ethanol. Publication No: DOE/GO/017057-1, Mascoma Corporation, 2012.
108. Domalski, E. and Hearing, E., NIST Chemistry WebBook Condensed Phase Heat Capacity Data, National Institute of Standards and Technology, Gaithersburg, MD, 2005.
109. Ryzak, T., Levin, D., B., Cicek, N. And Sparling, R., End-Product Induced Metabolic Shifts in *Clostridium thermocellum* ATCC 27405, Applied and Environmental Microbiology, 92,1 (2011) 199.

110. Li, H., F., Knutson, B., L., Nokes, S., E., Lynn, B., C. and Flythe, M., D., Metabolic Control of *Clostridium thermocellum* via Inhibition of Hydrogenase Activity and the Glucose Transport Rate, Applied and Environmental Microbiology, 93,4 (2012) 1777-1784.
111. Weimer, P. J. and Zeikus J.G., Fermentation of cellulose and cellobiose by *Clostridium thermocellum* in the absence of *Methanobacterium thermoautotrophicum*, Applied and Environmental Microbiology, 33,2 (1977) 289-297.
112. Brekasis, D. and Paget, M., S., A Novel Sensor of NADH/NAD<sup>+</sup> Redox Poise in *Streptomyces coelicolor* A3 (2), The EMBO Journal, 22,18 (2003) 4856-4865.
113. Green, J. and Paget, M., S., Bacterial redox sensors, Nature Reviews Microbiology, 2,12 (2004) 954-966.
114. Pagels, M., Fuchs, S., Pané-Farré, J., Kohler, C., Menschner, L., Hecker, M. And Lalk, M., Redox Sensing by A Rex-Family Repressor is Involved in the Regulation of Anaerobic Gene Expression in *Staphylococcus aureus*, Molecular Microbiology, 76,5 (2010) 1142-1161.
115. Wietzke, M. and Bahl, H., The Redox-Sensing Protein Rex, A Transcriptional Regulator of Solventogenesis in *Clostridium acetobutylicum*, Applied and Environmental Microbiology, 96,3 (2012), 749-761.
116. Hu, L., Huang, H., Yuan, H., Tao, F., Xie, H. and Wang, S., Rex in *Clostridium kluyveri* is A Global Redox-Sensing Transcriptional Regulator, Journal of Biotechnology, 233 (2016) 17-25.
117. Pei, J., Zhou, Q., Jing, Q., Li, L., Dai, C., Li, H. and Shao, W., The Mechanism for Regulating Ethanol Fermentation by Redox Levels in *Thermoanaerobacter ethanolicus*, Metabolic Engineering, 13,2 (2011) 186-193.
118. Taylor, B., L. and Zhulin, I., B., PAS domains: Internal Sensors of Oxygen, Redox Potential, and Light, Microbiology and Molecular Biology Reviews, 63,2 (1999) 479-506.
119. Tripathi, S., A., Olson, D., G., Argyros, D., A., Miller, B., B., Barrett, T., F., Murphy, D., M. and Hogsett, D., A., Development of *pyrF*-Based Genetic System for Targeted Gene Deletion in *Clostridium thermocellum* and Creation of A *pta* mutant, Applied and Environmental Microbiology, 76,19 (2010) 6591-6599.
120. Argyros, D., A., Tripathi, S., A., Barrett, T., F., Rogers, S., R., Feinberg, L., F., Olson, D., G. and Caiazza, N., C., High Ethanol Titers from Cellulose by Using Metabolically Engineered Thermophilic, Anaerobic Microbes, Applied and Environmental Microbiology, 77,23 (2011) 8288-8294.
121. Schwarz, W., The Cellulosome and Cellulose Degradation by Anaerobic Bacteria, Applied and Environmental Microbiology, 5,6 (2001) 634-649.

122. Ding, S., Y., Xu, Q., Crowley, M., Zeng, Y., Nimlos, M., Lamed, R. and Himmel, M., E., A Biophysical Perspective On the Cellulosome: New Opportunities for Biomass Conversion, Current Opinion in Biotechnology, 19,3 (2008) 218-227.
123. Ellis, L. D., Holwerda, E. K., Hogsett, D., Rogers, S., Shao, X., Tschaplinski, T. and Lynd, L., R., Closing the Carbon Balance for Fermentation by *Clostridium thermocellum* (ATCC 27405), Bioresource Technology, 103,1 (2012) 293-299.
124. Shao, X., Raman, B., Zhu, M., Mielenz, J., R., Brown, S., D., Guss, A., M. and Lynd, L., R., Mutant Selection and Phenotypic and Genetic Characterization of Ethanol-Tolerant Strains of *Clostridium thermocellum*, Applied and Environmental Microbiology, 92,3 (2011), 641-652.
125. Guss, A., M., Olson, D., G., Caiazza, N., C. and Lynd, L., R., Dcm Methylation is Detrimental to Plasmid Transformation in *Clostridium thermocellum*, Biotechnology for Biofuels, 5,1 (2012) 30.
126. Berríos-Rivera, S., J., Bennett, G., N. and San, K., Y., Metabolic Engineering of *Escherichia coli*: Increase of NADH Availability by Overexpressing an NAD<sup>+</sup>-Dependent Formate Dehydrogenase, Metabolic Engineering, 4,3 (2002) 217-229.
127. Van Zyl, L., J., Taylor, M., P., Eley, K., Tuffin, M. and Cowan, D., A., Engineering Pyruvate Decarboxylase-Mediated Ethanol Production in the Thermophilic host *Geobacillus thermoglucosidasius*, Applied and Environmental Microbiology, 98, 3 (2014) 1247-1259.
128. Feinberg, L., Foden, J., Barrett, T., Davenport, K., W., Bruce, D., Detter, C. and Cheng, J., F., Complete Genome Sequence of the Cellulolytic Thermophile *Clostridium thermocellum* DSM1313, Journal of Bacteriology, 193,11 (2011) 2906-2907.
129. Biswas, R., Zheng, T., Olson, D., G., Lynd, L., R. and Guss, A., M., Elimination of Hydrogenase Active Site Assembly Blocks H<sub>2</sub> Production and Increases Ethanol Yield in *Clostridium thermocellum*, Biotechnology for Biofuels, 8,1 (2015) 20.
130. Strobel, H., P. and Lynn, B., Proteomic Analysis of Ethanol Sensitivity in *Clostridium thermocellum*, Proceeding of the National Meeting of the American Society of Microbiology, May 2004, 0-094
131. Adams, M., W., Mortenson, L. E. and Chen, J., S., Hydrogenase, Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics, 594,2-3 (1980) 105-176.
132. Jugder, B., E., Chen, Z., Ping, D., T., T., Lebhar, H., Welch, J. and Marquis, C., P., An Analysis of the Changes in Soluble Hydrogenase and Global Gene Expression in *Cupriavidus necator* (*Ralstonia eutropha*) H16 Grown in Heterotrophic Diauxic Batch Culture, Microbial Cell Factories, 14,1 (2015) 42.
133. Mayhew, S., G. and O'Connor M., E., Structure and Mechanism of Bacterial Hydrogenase, Trends in Biochemical Sciences, 7,1 (1982) 18-21.

134. Vignais, Paulette M., and Bernard Billoud. "Occurrence, classification, and biological function of hydrogenases: an overview." *Chemical reviews* 107.10 (2007): 4206-4272.
135. Vignais, P., M. and Colbeau., A., Molecular Biology of Microbial Hydrogenases, Current Issues in Molecular Biology, 6,2 (2004) 159-188.
136. Peters, J., W., Structure and Mechanism of Iron-Only Hydrogenases, Current Opinion in Structural Biology, 9,6 (1999) 670-676.
137. Jugder, B., E., Welch, J., Aguey-Zinsou, K., F. and Marquis, C., P., Fundamentals and Electrochemical Applications of [Ni-Fe]-Uptake Hydrogenases, RSC Advances, 3,22 (2013) 8142-8159.
138. Winkler, M., Esselborn, J. and Happe, T., Molecular Basis of [FeFe]-Hydrogenase Function: An Insight Into the Complex Interplay Between Protein and Catalytic Cofactor, Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1827,8 (2013) 974-985.
139. Koç, E. and Şenel, M., C., Dünyada ve Türkiye’de Enerji Durumu-Genel Değerlendirme, Mühendis ve Makina, 54,639 (2013) 32-44.
140. Ünal, N., K., Türkiye ve Avrupa Birliği’nde Biyoyakıt. Türk Tarım ve Doğa Bilimleri, 2,1 (2015) 26–33.
141. <http://www.tepge.gov.tr/upload/attachments/Biyoyakit204.pdf> Tarımsal Ekonomi ve Politika Geliştirme Enstitüsü, Dünya ve Türkiye Biyo-enerji Piyasalarındaki Gelişmelerin ve Potansiyel Değişikliklerin Türk Tarım ve Hayvancılık Sektörleri Üzerindeki Etkilerinin Modellenmesi ve Türkiye için Biyo-enerji Politika Alternatiflerinin Oluşturulması 20 Kasım 2017.
142. Karaosmanoğlu, F., Türkiye Biyoyakıt Potansiyeli ve Son Gelişmeler. [http://www.dektmk.org.tr/pdf/enerji\\_kongresi\\_10/filizkaraosmanoglu.pdf](http://www.dektmk.org.tr/pdf/enerji_kongresi_10/filizkaraosmanoglu.pdf) 20 Aralık 2017.
143. Mai, V., Lorenz, W., W. and Wiegel, J., Transformation of *Thermoanaerobacterium sp.* Strain JW/SL-YS485 with Plasmid pIKM1 Conferring Kanamycin Resistance, FEMS Microbiology Letters, 148, 2 (1997) 163-167.
144. Hon, S., Olson, D., G., Holwerda, E., K., Lanahan, A., A., Murphy, S., J., Maloney, M., I. and Lynd, L. R. The ethanol pathway from *Thermoanaerobacterium saccharolyticum* improves ethanol production in *Clostridium thermocellum*, Metabolic Engineering, 42 (2017), 175-184.
145. Hon, S., Lanahan, A., A., Tian, L., Giannone, R., J., Hettich, R., L., Olson, D., G. and Lynd, L., R., Development of A Plasmid-Based Expression System in *Clostridium thermocellum* and Its Use to Screen Heterologous Expression of Bifunctional Alcohol Dehydrogenases (*adhEs*), Metabolic Engineering Communications, 3 (2016) 120-129.

146. Gibson, D., G., Enzymatic Assembly of Overlapping DNA Fragments, Methods in Enzymology, 498, (2011) 349-361.
147. Sambrook, J. and Russell, D., W., Molecular cloning: A laboratory Manual, Third Edition, Cold Spring Harbor Laboratory Press, New York, 2001.
148. Zhou, J., Olson, D., G., Lanahan, A., Tian, L., Murphy, S., J. L., Lo, J. and Lynd, L., R. Physiological Roles of Pyruvate Ferredoxin Oxidoreductase and Pyruvate Formate-Lyase in *Thermoanaerobacterium saccharolyticum* JW/SL-YS485, Biotechnology for Biofuels, 8,1 (2015) 138.
149. William, S., Feil, H. and Copeland, A., Bacterial Genomic DNA Isolation Using CTAB, Sigma, 50 (2012) 6876.
150. Altschul, S., F., Gish, W., Miller, W., Myers, E., W. and Lipman, D., J., Basic Local Alignment Search Tool, Journal of Molecular Biology, 215 (1990) 403-410.
151. Eminoğlu, A., Murphy, S., J., L., Maloney, M., Lanahan, A., Giannone, R., J., Hettich, R., L., Tripathi, S., A., Beldüz, A., O., Lynd L., R. and Olson, D., G., Deletion of the *hfsB* Gene Increases Ethanol Production in *Thermoanaerobacterium saccharolyticum* and Several Other Thermophilic Anaerobic Bacteria, Biotechnology for Biofuels, 10,1 (2017) 282.
152. Cui, X. and Churchill, G., A., Statistical Tests for Differential Expression in cDNA Microarray Experiments, Genome Biology, 4,4 (2003) 210.
153. Giannone, R., J., Huber, H., Karpinets, T., Heimerl, T., Küper, U., Rachel, R. And Podar, M., Proteomic Characterization of Cellular and Molecular Processes that Enable the *Nanoarchaeum equitans*-*Ignicoccus hospitalis* Relationship, PloS One, 6,8 (2011) e22942.
154. Giannone, R., J., Wurch, L., L., Heimerl, T., Martin, S., Yang, Z., Huber, H. and Podar, M., Life on the edge: functional genomic response of *Ignicoccus hospitalis* to the presence of *Nanoarchaeum equitans*, ISME Journal, 9,1 (2015) 101-14.
155. Tabb, D., L., Fernando, C., G. and Chambers M., C., MyriMatch: Highly Accurate Tandem Mass Spectral Peptide Identification by Multivariate Hypergeometric Analysis, Journal of Proteome Research, 6,2 (2008) 654-661.
156. Ma, Z., Q., Dasari, S., Chambers, M., C., Litton, M., D., Sobecki, S., M., Zimmerman, L., J. and Tabb, D., L., IDPicker 2.0: Improved Protein Assembly with High Discrimination Peptide Identification Filtering, Journal of Proteome Research, 8,8 (2009) 3872-3881.
157. Hogsett, D., Cellulose Hydrolysis and Fermentation by *Clostridium thermocellum* for the Production of Ethanol, PhD thesis, Thayer School of Engineering, Hanover, 1995.
158. Lo, J., Zheng, T., Hon, S., Olson, D., G. and Lynd, L., R., The Bifunctional Alcohol and Aldehyde dehydrogenase Gene, *adhE*, is Necessary for Ethanol Production in

- Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*, Journal of Bacteriology, 197, 8 (2015) 1386-1393.
159. Ma, K. and Adams, M. W., Ferredoxin:NADP Oxidoreductase from *Pyrococcus furiosus*, Methods in Enzymology, 334 (2001) 40–45.
160. Zheng, T., Olson, D., G., Murphy, S., J., Shao, X., Tian, L. and Lynd, L., R., Both *adhE* and A Separate NADPH-Dependent Alcohol dehydrogenase Gene, *adhA*, are Necessary for High Ethanol Production in *Thermoanaerobacterium saccharolyticum*, Journal of Bacteriology, 199,3 (2017) e00542-16.
161. Zheng, T., Olson, D., G., Tian, L., Bomble, Y., J., Himmel, M., E., Lo, J. and Lynd, L. R., Cofactor Specificity of the Bifunctional Alcohol and Aldehyde Dehydrogenase (*AdhE*) in Wild-Type and Mutant *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*, Journal of Bacteriology, 197,15 (2015) 2610-2619.
162. Zhang, L., Nie, X., Ravcheev, D., A., Rodionov, D., A., Sheng, J., Gu, Y. and Yang, C., Redox-Responsive Repressor Rex Modulates Alcohol Production and Oxidative Stress Tolerance in *Clostridium acetobutylicum*, Journal of Bacteriology, 196,22 (2014) 3949-3963.
163. Finn, R. D., Coggill, P., Eberhardt, R., Y., Eddy, S., R., Mistry, J., Mitchell, A., L. and Salazar, G., A., The Pfam Protein Families Database: Towards A More Sustainable Future, Nucleic Acids Research, 44,1 (2016) 279-285.



## CURRICULUM VITAE

Ayşenur EMİNOĞLU was born in Konya, Turkey in 1979. She graduated from the Atatürk Girl High School in 1996. She received her B.Sc. from the Selçuk University, Konya in 2003 at second place. After graduation, she worked as a high school teacher, in Biology from 2004 to 2009. After completing her master's degree in Molecular Biology at Recep Tayyip Erdogan University in 2012, she started her Ph.D. education at KTU, Department of Molecular Biology, in 2013. From 2015 to 2017 she worked as a visiting researcher at Department of Chemical and Biochemical Engineering at Thayer School of Engineering at Dartmouth, USA with the support of TÜBİTAK BİDEP (2214/A). Since 2011, she has been working at Recep Tayyip Erdogan University, Rize as a Research Assistant.

### Publication's From This Thesis;

#### Paper:

**Eminoğlu, A.** , Murphy, S., J. , Maloney, M. , Lanahan, A. , Giannone, R., J. , Hettich R., L., Tripathi S., A., Beldüz A., O., Lynd L., R. and Olson D.G., Deletion of the *hfsB* gene increases ethanol production in *Thermoanaerobacterium saccharolyticum* and several other thermophilic anaerobic bacteria, Biotechnology for Biofuels, 282, (2017) 1-11