Metabolic engineering yeast cells for medium-chained biofuel synthesis

Li, Xiang

2015

Li, X. (2015). Metabolic engineering yeast cells for medium-chained biofuel synthesis.
Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/62904>

<https://doi.org/10.32657/10356/62904>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**METABOLIC ENGINEERING YEAST CELLS FOR
MEDIUM-CHAINED BIOFUEL SYNTHESIS**

**LI XIANG
SCHOOL OF CHEMICAL AND BIOMEDICAL ENGINEERING
2015**

METABOLIC ENGINEERING YEAST CELLS FOR MEDIUM-CHAINED BIOFUEL SYNTHESIS

LI XIANG

LI XIANG

School of Chemical and Biomedical Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of

Doctor of Philosophy

2015

ACKNOWLEDGEMENTS

Four years at Nanyang Technological University were unforgettable. NTU provided me a platform to pursue my degree and opportunity to get to know the most frontier areas and so many outstanding researchers who generously offered me sincere help. Herein, I would like to extend my deep appreciations to them.

First and foremost, I would like to thank my supervisor, Prof. Chen Wei Ning William. He kindly provided me the opportunity to pursue my PhD, which has been a great honor. He has always been so kind and supportive and the family-like atmosphere in our research group really has made the four years enjoyable. His professional guidance, enlightening instructions and patient supervisions supported me in every stage in the process of generating this essay.

I shall express my appreciations to my colleagues: Dr. Feng Huixing, Dr. Zhang Jianhua, Dr. Wang Mingxuan, Dr. Bai Jing, Dr. Zhou Yusi, Dr. Laleh Sadrolodabae, Dr. Tan Yi Lin Jane, Mr. Tan Kee Yang, Ms. Tang Xiaoling, Ms. Shi Jiahua, Ms. Chen Liwei, Ms. Zhao Guili, Ms. Lee Jie Lin Jaslyn and Ms. Toh Shi Hui, for their constructive suggestions and kind encouragement, for their support and friendship.

Then I shall express my highly grateful to the professional officer Dr Fang Ning, Dr Ong Teng Teng, Dr. Wang Xiujuan, Ms. Tan Lay Yian and Ms. Gan Chew Mei Jessica, for their professional technician support.

Last but not least, I would like to express my love to my parents and my fiancé, to whom this dissertation is dedicated to. Their unconditional love and trust have supported me and would always be my greatest wealth.

Finally, I appreciate the financial support from Competitive Research Program that funded the research in this dissertation.

CONTENTS

LIST OF FIGURES	i
LIST OF TABLES	ix
ABBREVIATIONS	xi
ABSTRACT.....	xv
CHAPTER 1	1
INTRODUCTION	1
1.1 Background	1
1.1.1 Development of biofuel	1
1.1.2 Renewable jet fuel.....	2
1.2 Biofuel production by metabolic engineering.....	4
1.3 Fatty acids metabolism in <i>S. cerevisiae</i>	5
1.4 Proteomics.....	7
1.4.1 Introductions on proteomics	7
1.4.2 Basic introduction of LC-MS.....	8
1.4.3 Proteomics basing on LC-MS	9
1.4.4 Stable-isotope technique	15
1.5 Basic introduction of GC	20
CHAPTER 2	23
OBJECTIVES	23
CHAPTER 3	25
RESEARCH DESIGN AND METHODS	25
3.1 Strains and media.....	25
3.2 Double deletion strain construction	25

3.3 Cloning target genes	32
3.3.1 LOX and HPL cloning	32
3.3.2 ADC cloning	34
3.4 Construction of functional <i>S. cerevisiae</i> strains	35
3.5 Growth curve test	37
3.6 Protein extraction and Western-blot analysis	37
3.6.1 Gel electrophoresis	38
3.6.2 Membrane transfer	41
3.6.3 Blocking	42
3.6.4 Detection	42
3.7 Biotransformation and identification	43
3.7.1 GC-FID	43
3.7.2 GC-MS	44
3.8 Proteomics	44
3.9 Protein identification and data analysis	46
3.10 Protein quantification and data analysis	48
CHAPTER 4	51
METABOLICALLY ENGINEERED YEAST CELLS AND MEDIUM- CHAINED HYDROCARBON BIOFUEL PRECURSORS SYNTHESIS	51
4.1 Introduction	51
4.1.1 The lipoxygenase pathway	51
4.1.2 Lipoxygenase	54
4.1.3 Hydroperoxide lyase	55
4.2 Experiment procedure	56
4.3 Results	56

4.3.1 Construction of recombinant plasmid 9LHP	56
4.3.2 Construction of double deletion strain	57
4.3.3 Construction of functional strains	60
4.3.4 Growth curve test	62
4.3.5 Western blot	63
4.3.6 Proteomics.....	64
4.3.7 GC-FID detection.....	78
4.4 Conclusions.....	82
4.5 Future directions	84
CHAPTER 5	87
METABOLICALLY ENGINEERED YEAST CELLS AND MEDIUM- CHAINED HYDROCARBON BIOFUEL SYNTHESIS (PRELIMINARY)	87
5.1 Introduction.....	87
5.2 Experiment procedure	90
5.3 Results and discussions.....	90
5.3.1 Construction of recombinant plasmid.....	90
5.3.2 Construction of functional strains	91
5.3.3 Growth curve test	92
5.3.4 SDS-PAGE and Western blot	93
5.3.5 GC-MS results	94
5.4 Conclusions.....	95
5.5 Future directions	99
REFERENCE.....	103
SUPPLEMENT.....	117
APPENDIX.....	133

LIST OF FIGURES

Chapter 1

- Fig. 1.1 Scheme of β -oxidation in yeast peroxisomes (23).** The protein complex, Pxa1p-Pxa2p (Pat1p-Pat2p) embedded in the peroxisomal membrane functions as a transporter, translocating activated FAs into peroxisomes for degradation. See descriptions on Page 6. 7
- Fig. 1.2 Mass spectrometers used in proteome research (29).** The mechanisms of ionization by ESI and MALDI are shown in upper left and upper right. Fig. 1.2a-f shows the configurations of specific mass spectrometers. Fig. 1.2a: reflector TOF instrument. Fig. 1.2b: TOF-TOF instrument. Fig. 1.2c: triple quadrupole or linear ion trap. Fig. 1.2d: quadrupole-TOF. Fig. 1.2e: 3-D ion trap. Fig. 1.2f: FT-MS instrument. See descriptions on Page 10-11. 12
- Fig. 1.3 Generic MS-based proteomics experiment (29).** Stage 1: target proteins extraction. Stage 2: digest proteins to linear peptides. Stage 3: peptides chromatography, ionization and elution into MS. Stage 4, protonated peptides enters the MS and a mass spectrum would be generated. Stage 5: data analysis. See descriptions on Page 13. 14
- Fig. 1.4 Scheme of stable-isotope protein labelling for quantitative proteomics (29).** Fig. 1.4a: Labeling proteins metabolically. Fig. 1.4b: Isotope tagging by chemical reaction. Fig. 1.4c: Stable-isotope incorporation via enzyme reaction. See descriptions on Page 15-16. 17
- Fig. 1.5 Principles of isobaric tagging by iTRAQ (35).** Fig. 1.5A: the diagram of the structures of the multiplexed isobaric tags. Fig. 1.5B: the

design of isotopes ^{13}C , ^{15}N , and ^{18}O . Fig. 1.5C: illustration of the isotopic tagging used to arrive at four isobaric combinations with four different reporter group masses.	18
---	----

Fig. 1.6 Scheme of the FID. FID works only for organic compounds. Two electrodes placed adjacent to a flame fueled by hydrogen/air near the exit of the column. Flow through of the organic compounds pyrolyzed by the flame would change the current or voltage between the electrodes, which were recorded as signals for translation into retention time. See descriptions on Page 20.	21
---	----

Fig. 1.7 The scheme of typical GC-MS. See descriptions on Page 21.....	22
---	----

Chapter 2

Fig. 2.1 Pathway of synthesis of medium-chained biofuel precursors.	24
--	----

Chapter 3

Fig. 3.1 Outline of the one-step gene disruption approach for generation of $\Delta pxa1\&2$ strain. See descriptions on Page 26.....	29
---	----

Fig. 3.2 Scheme of recombinant plasmid 9LOX-9HPL-pESC (9LHP), 11820 kb. See descriptions on Page 32-33.....	33
--	----

Fig. 3.3 Scheme of recombinant plasmid 9LOX-9HPL-ADC-pESC (9LHPA), 13117 bp. See descriptions on Page 33-34.	35
--	----

Fig. 3.4 Flowchart of iTRAQ-based quantitative proteomics experimental design. See descriptions on Page 44-45.	46
--	----

Fig. 3.5 Workflow of on-line 2D nano-LC. (a) The peptides eluted through SCX column and trapped onto enrichment column. (b) HPLC-chip on	
---	--

analytical mode, the previously trapped peptides were eluted and	
analyzed by analytical column. See descriptions on Page 46-47.....	48

Chapter 4

Fig. 4.1 Diagram of short-chained aldehyde forming system in plants (51).

See description on Page 52.	53
----------------------------------	----

Fig. 4.2 Combined results of four different double digestions of recombinant vector 9LOX-9HPL-pESC (11820 bp) for confirmation.

a) 5 k and 7 k with <i>SacI</i> (3320) and <i>Sall</i> (8160); b) 5 k and 7 k with <i>PacI</i> (3312) and <i>Sall</i> (8160); c) 5 k and 7 k with <i>SacI</i> (3320) and <i>XhoI</i> (8196); d) 5 k and 7 k with <i>PacI</i> (3312) and <i>NheI</i> (8218). (GeneRuler 1 kb Plus DNA Ladder, ready-to-use, Fermentas).....	57
--	----

Fig. 4.3 The gel electrophoresis of gene disruption confirmation. Lane 1:

A+His B, product size of 771 bp. Lane 2: A+B, band. Lane 3:C+D, no band . Lane 4:His C+D, product size of 468. See descriptions on Page 57.	58
--	----

Fig. 4.4 The gel electrophoresis of deletion strain. Lane 1: A1+His B1,

product size of 1205 bp. Lane 2: A1+B1, no band. Lane 3:A2+His B2, product size of 1456 bp. Lane 4, A2+B2, no band. Lane 5:A3+His B3, product size of 1193. Lane 6, A3+B3, no band. See descriptions on Page 58.....	59
--	----

Fig. 4.5 The gel electrophoresis f deletion strain. Lane 1/2: A1+D, product

size of 2367 bp. Lane 3/4: blank. Lane 5/6: A2+D, product size of 2359. Lane 7:A3+D, product size of 1863. See descriptions on Page 58.....	59
---	----

- Fig. 4.6 Colony PCR results of the functional strains.** Left: LOX, size=1145 bp; Right: HPL, size= 730 bp. Lane 1: WT-9LHP. Lane 2: $\Delta pxa1$ -9LHP. Lane 3: $\Delta pxa1$ -9LHP. Lane 4: $\Delta pxa1 \& 2$ -9LHP. See descriptions on Page 60. 61
- Fig. 4.7 Colony PCR results of the control strains against pESC, size= 1323 bp.** Lane 1: WT-pESC. Lane 2: $\Delta pxa1$ - pESC. Lane 3: $\Delta pxa1$ -pESC. Lane 4: $\Delta pxa1 \& 2$ - pESC. See descriptions on Page 60. 61
- Fig. 4.8 Growth curve of *S. cerevisiae* strains WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP, $\Delta pxa1 \& 2$ -9LHP and WT-pESC, $\Delta pxa1$ -pESC, $\Delta pxa2$ -pESC, $\Delta pxa1 \& 2$ -pESC as controls.** 62
- Fig. 4.9 Western blot of target genes in *S.cerevisiae* control strains and functional strains:** a) 9LOX, 99kDa; b) 9HPL, 55kDa. Lane 1: WT-pESC. Lane 2: $\Delta pxa1$ -pESC. Lane 3: $\Delta pxa2$ -pESC. Lane 4: $\Delta pxa1 \& 2$ -pESC. Lane 5: WT-9LHP. Lane 6: $\Delta pxa1$ -9LHP. Lane 7: $\Delta pxa2$ -9LHP. Lane 8: $\Delta pxa1 \& 2$ -9LHP. 63
- Fig. 4.10 Representative peptide fragmentation spectrum of glucose-6-phosphate isomerase: (R)AVYHVALR(N) in WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1 \& 2$ -9LHP combined sample.** More details in Fig. S2, Page 121..... 64
- Fig. 4.11 Western blot results of phosphoglycerate kinase 1 (45 kDa) and glyceraldehyde-3-phosphate dehydrogenase (37 kDa) (A) Protein expression changes. (B) Quantification of protein level changes based on Western blot analysis.** 70

Fig. 4.12 Diagram of Leloir pathway (69). Galactokinase and galactose-1-phosphate uridylyltransferase catalyzing 2 irreversible reactions were identified in proteomics study. See descriptions on Page 71.....	75
Fig. 4.13 Diagram of glycolysis (69). <i>hexokinase-1</i> , glucose-6-phosphate isomerase, <i>phosphofructokinase</i> , fructose-bisphosphate aldolase, triosephosphate isomerase, glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase, phosphoglycerate mutase 1, enolase and <i>pyruvate kinase</i> were identified in proteomics (Enzymes catalyzing rate-limiting steps are in italic). See descriptions on Page 71-72.	76
Fig. 4.14 Diagram of TCA cycle (69). <i>Citrate synthase</i> , mitochondrial and aconitase, mitochondrial (Enzymes catalyzing rate-limiting steps are in italic). See descriptions on Page 72.	77
Fig. 4.15 GC spectra: blue: 3(Z)-nonenal standard; red: $\Delta pxa1$ & 2-9LHP; green: $\Delta pxa1$ & 2-pESC. Retention time at 8.82 min was identified at 3(Z)-nonenal. See descriptions on Page 78.....	79
Fig. 4.16 3(Z)-nonenal production levels by constructed strains. See descriptions on Page 78.	80

Chapter 5

Fig. 5.1 Overall approaches to biosynthesize medium-chained biofuel.	89
Fig. 5.2 Two different double digestion combination results of recombinant plasmid 9LHPA(13117bp) for confirmation: a) 5k and 8k with <i>SacI</i> (4627) and <i>NheI</i> (9525); b) 8 k and 5 k with <i>ClaI</i> (1604) and <i>NheI</i> (9525). (GeneRuler 1 kb Plus DNA Ladder, ready-to-use, Fermentas). See descriptions on Page 90.....	91

Fig. 5.3 Colony PCR results of the functional strains against ADC gene, size=460 bp. Lane 4: WT-9LHPA. Lane 5: $\Delta pxa1$ -9LHPA. Lane 6: $\Delta pxa1$ -9LHPA. Lane 7: $\Delta pxa1$ & 2-9LHPA. See descriptions on Page 91- 92.....	92
Fig. 5.4 Growth curve tested of functional strains WT-9LHPA, $\Delta pxa1$- 9LHPA, $\Delta pxa2$-9LHPA and $\Delta pxa1$ & 2-9LHPA. See descriptions on Page 92-93.	93
Fig. 5.5 Western blot of ADC (28 kDa) in <i>S.cerevisiae</i> control strains and functional strains; Lane 1: WT-pESC. Lane 2: $\Delta pxa1$ -pESC. Lane 3: $\Delta pxa2$ -pESC. Lane 4: $\Delta pxa1$ & 2-pESC. Lane 5: WT-9LHPA. Lane 6: $\Delta pxa1$ -9LHPA. Lane 7: $\Delta pxa2$ -9LHPA. Lane 8: $\Delta pxa1$ & 2-9LHPA. See descriptions on Page 93-94.	94
Fig. 5.6 GC spectra. Peaks at retention time of 8.10 min were identified to be 3(Z)-nonenal. Red, WT-LHPA; Green, YKL-LHPA; Black, YPD- LHPA; Blue: $\Delta pxa1$ & 2-LHPA. See descriptions on Page 94-95.	96
Fig. 5.7 GC spectra. Peaks at retention time of 8.10 min were identified to be 3(Z)-nonenal. Red, WT-LHP; Green, YKL-LHP; Black, YPD-LHP; Blue: $\Delta pxa1$ & 2-LHP. See descriptions on Page 94-95.	97
Fig. 5.8 MS spectrum of 3(Z)-nonenal: The peak at retention time of 8.10 min. See descriptions on Page 94-95.	98

Supplement

Fig. S1 Total intensity chromatogram results of peptides eluted by gradient concentrations of APS. See descriptions on Page 64.	120
---	-----

Fig. S2 LC-MS qualification results of representative peptide fragmentation spectrum of glucose-6-phosphate isomerase. See descriptions on Page 64.	121
Fig. S3 Peptide summary of glucose-6-phosphate isomerase. See descriptions on Page 64.	122
Fig. S4 MS digest results of glucose-6-phosphate isomerase (part 1). See descriptions on Page 64.	123
Fig. S5 MS digest results of glucose-6-phosphate isomerase (part 2). See descriptions on Page 64.	124
Fig. S6 MS digest results of glucose-6-phosphate isomerase (part 3). See descriptions on Page 64.	125
Fig. S7 MS digest results of glucose-6-phosphate isomerase (part 4). See descriptions on Page 64.	126
Fig. S8 MS digest results of glucose-6-phosphate isomerase (part 5). See descriptions on Page 64.	127

LIST OF TABLES

Chapter 3

Table 3.1 Primers, plasmids and strains used into study	27
Table 3.2 Primer designs to confirm gene disruption.....	30
Table 3.3 The combinations of primers used in PCR and resulting product sizes.	31
Table 3.4 Primers for colony PCR.....	37
Table 3.5 SDS-PAGE protocol of separating gel.....	39
Table 3.6 SDS-PAGE protocol of stacking gel	40

Chapter 4

Table 4.1 Relative changes in protein expression of <i>S. cerevisiae</i> WT- 9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1\&2$ -9LHP	67
Table 4.2 Production of functional strains and carbon recovery rates.....	81

Supplement

Table S1 Data of growth curve Fig. 4.8.....	117
Table S2 Heat map of proteomics results in Table 4.1.	128
Table S3 Data of growth curve Fig. 5.4.....	131

ABBREVIATIONS

ABC	ATP binding cassette
ADC	Aldehyde decarbonylase
ADO	Aldehyde-deformylating oxygenase
AHR	Aldehyde reductase
AOS	Allene oxide synthase
APS	Ammonium persulfate
ATP	Adenosine triphosphate
Bio-SPK	Bio-derived synthetic paraffinic kerosene
BSA	Bovine serum albumin
CAR	Carboxylic acid reductase
CID	Collision-induced dissociation
DES	Divinyl ether synthase
ECL	Enhanced chemi-luminescence
EI	Electron ionization
EPR	Electron paramagnetic resonance
EUROSCARF	European <i>Saccharomyces cerevisiae</i> Archive for Functional analysis
ESI	Electrospray ionization

FA	Fatty acid
FID	Flame ionization detector
FT-ICR	Fourier transform ion cyclotron resonance
FT-SPK	Fischer-Tropsch Synthetic Paraffinic Kerosene
GC	Gas chromatography
GC-FID	Gas chromatography-flame ionization detector
GC-MS	Gas chromatography-mass spectrometry
G6P	Glucose-6-phosphate
HPL	Hydroperoxide lyase
HPOD	Hydroperoxide
9-HPOD	Linoleic acid 9S-hydroperoxide
13-HPOD	Linoleic acid 13S-hydroperoxide
9-HPOT	α -linolenic acid 9S-hydroperoxide
13-HPOT	α -linolenic acid 13S-hydroperoxide
HRP	Horseradish peroxidase
ICP	Inductively coupled plasma
iTRAQ	Isobaric tags for relative and absolute quantitation
LC	Liquid chromatography
LCFA	Long-chained fatty acid
LC-MS	Liquid chromatography mass spectrometry

LC-MS/MS	Liquid chromatography with tandem mass spectrometry
LOX	Lipoxygenase
MALDI	Matrix-assisted laser desorption/ionization
MCFA	Medium-chained fatty acid
MS	Mass spectrometry
OD	Optical density
PBS	Phosphate buffered saline
PBST	Phosphate buffered saline-tween 20
PCR	Polymerase chain reaction
PUFA	Polyunsaturated fatty acid
PVDF	Polyvinylidene fluoride
Q-TOF	Quadrupole time-of-flight
SCX	Strong cation-exchange
SDS	Sodium dodecyl sulphate
SDS-PAGE	sulphate-poly acrylamide gel electrophoresis
SIM	Selected ion monitoring
TAG	Triacylglycerol
TCA cycle	Tricarboxylic acid cycle
TEMED	N,N,N',N'-Tetramethylethylenediamine
VLCFA	Very long-chained fatty acid

ABSTRACT

Biotechnologies could potentially support renewable fuels production by modifying existing pathways or by creating synthetic routes via metabolic engineering. Recently, metabolic engineering technique has been applied to synthesize biofuels to meet the exploding energy demand and petroleum shortage.

In this work, *Saccharomyces cerevisiae* capable of synthesizing medium-chained hydrocarbons was constructed. The hydroperoxide pathway from almond which consists of lipoxygenase (LOX) and hydroperoxide lyase (HPL), catalyzed linoleic acid to 3(Z)-nonenal as medium-chained biofuel precursor, was introduced into *S. cerevisiae* wild type, $\Delta pxa1$, $\Delta pxa2$ and $\Delta pxa1\&2$ mutants to construct whole-cell based biocatalysts. Then aldehyde decarbonylase (ADC) was introduced to catalyze the 3(Z)-nonenal into octene, which was the final target hydrocarbon.

Previous studies have discovered that β -oxidation in *S. cerevisiae* was confined in peroxisomes. Long-chained fatty acids (LCFAs) were first activated in cytoplasm and then transported into peroxisomes for degradation by the ATP binding cassette (ABC) transporter Pxa1/Pxa2. Therefore, the single mutants and double mutant of this transporter were adopted, aiming to retain the absorbed LCFAs in cytoplasm.

Proteomics study through LC-MS/MS approach was carried out to implement the investigation of overall protein levels. 31 proteins showed different expression levels among the functional strains WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP, $\Delta pxa1\&2$ -9LHP. The proteins involved in galactose metabolism, glycolysis, tricarboxylic acid cycle (TCA cycle) and ATP synthesis were notably up-regulated in the strain $\Delta pxa1\&2$ -9LHP, which suggested the higher activities of metabolism

and energy provision. Furthermore, several proteins involved in amino-acid metabolism and protein biosynthesis, which would support the expression of the exogenous genes, were also significantly up-regulated in the same strain. The proteomics study indicated that functional strain $\Delta pxa1 \& 2$ -9LHP may display the highest biotransformation efficiency.

After the proteomics study, the biotransformation capabilities of the functional strains were determined. Biotransformation using resting cells with linoleic acid added to the culture media was carried out and produced 3(Z)-nonenal was qualified and quantified with gas chromatography-flame ionization detector (GC-FID). The functional $\Delta pxa1 \& 2$ -9LHP strain showed the highest yield of up to 1.21 mg/L. The carbon recovery rate was calculated to be 12.1%. The biotransformation results corresponded to our expectations from the proteomics study. This indicated that double disruption of the peroxisomal transporter would influence the flux of absorbed linoleic acid and further help to retain the absorbed linoleic acid in cytoplasm for degradation through the hydroperoxide pathway.

ADC introduced into the functional strains was successfully expressed; however, gas chromatography-mass spectrometry (GC-MS) detection proved that the level of the produced octene was non-detectable.

In the future, the overall optimization should be carried out, including: 1) the selection of microbial host and culture conditions, changing the expression host into other oleaginous species, *Yarrowia lipolytica* for instance. 2) the expression conditions of ADC, including structure-based modification to improved enzyme properties, the steric configuration and the kinetics properties. 3) cofactors and

byproducts balancing, to make the metabolically engineered microbial host reach a novel systematic balance.

The approach described here would potentially contribute to the production of medium-chained biofuel precursors and then biofuels, which is one step further towards the goal of low-cost renewable transportation fuels.

CHAPTER 1

INTRODUCTION

1.1 Background

Fossil fuels, including petroleum, coal oil and natural gas, have supported the world's rapid development in the past century. The energy crisis in early 1970s triggered awareness and urged the development and utilization of renewable energies such as solar energy, wind energy, ocean energy and geothermal energy. Among them, biofuels, as an emerging technology, opened up new possibilities (1-3).

1.1.1 Development of biofuel

Biofuels are renewable fuels produced from solar energy which is bio-chemically stored as high-energy organic compounds by organisms. CO₂ in atmosphere is first absorbed and fixed in sugars by photosynthesis. Then the sugars can be converted into fuels by metabolically engineered microorganisms. Combustion of the biofuels re-releases CO₂ back into the atmosphere. This closed CO₂ cycle is driven by the energy of the sun and constitutes a more carbon-neutral process than the direct burning of fossil fuels, the cause of ever-increasing CO₂ levels.

The 1st generation or conventional biofuels are made from food crops, such as sugar, starch, or vegetable oil. Mature technology and commercial market for the production and usage are in place. Nowadays, the 1st generation biofuel ethanol, produced from microbial fermentation of starch from feedstock is present in over

90% of U.S. gasoline pumps, at up to 10% concentrations (3). However, the sustainable and economic production of 1st generation biofuels has raised some problems to solve, such as high production and processing costs and up-regulating the price of foodstuff. This situation has stimulated the development of 2nd generation biofuels produced from non-food biomass. The non-food biomass included lingo-cellulosic biomass or woody crops, agricultural residues and non-edible parts of plants.

2nd generation biofuels could help to solve the problems faced by 1st generation biofuels and also would support a larger proportion of global fuel supply more environmental-friendly. The problem that second generation biofuel encountered is the difficulty of extracting the useful feedstock from cellulosic material and the technology for converting them into biofuels is still in an early stage of development.

1.1.2 Renewable jet fuel

Jet fuels are different from traditional engine fuels. Their physicochemical properties are specially designed to fit the engine-specific requirements and operational conditions of aircrafts. The content and proportions of hydrocarbon constituents in jet fuels are crucial determinants including the energy content, combustibility, density and fluidity.

Jet fuel traditionally corresponds to the kerosene distillation fraction of petroleum. Currently, the renewable alternative jet fuels are provided by Bio derived synthetic paraffinic kerosene (Bio-SPK) and Fischer-Tropsch synthetic paraffinic kerosene (FT-SPK) (4).

Bio-SPK is produced by transesterification of triacylglycerols and FAs extracted from plants, algae or recycled sources. Then alkanes with desired properties

including chain length, saturation level and branching are produced by hydrocracking and hydroprocessing.

FT-SPK is obtained through pyrolysis of biomass into synthetic gas, Fischer-Tropsch synthesis of longer chain alkanes, hydro-processing and separation. Though the chemical compositions of SPKs and petroleum-based fuels are clearly different, they are still very similar in their key technical properties and performance in modern jet aircrafts.

Alkanes are the major components of jet fuels. The production of similar hydrocarbons through microbial biosynthesis pathways is promising as an alternative for biofuel production (4, 5). An important future goal is to simplify the fuel production process, and reduce the number of steps by directly generating the desired fuel from sunlight and CO₂, instead of proceeding through terrestrial biomass. For this, photosynthetic aquatic microorganisms such as cyanobacteria and eukaryotic algae provide an attractive platform for the direct conversion of solar energy into engine-ready fuels that preferentially are excreted from the host.

The current production efficiency cannot provide the possibility of industrialization. However, with more devoted efforts, we look forward to a future platform for the direct, environmentally beneficial and economic production of renewable fuels from biomass.

This research mainly focuses on the biosynthesis of medium-chained hydrocarbon biofuels and precursor.

1.2 Biofuel production by metabolic engineering

Biofuels, such as ethanol, biodiesel and hydrocarbons, were defined as a variety of fuels produced from biomass resources, basing on metabolic engineering microorganisms (6-9). Various potential biofuels has been produced while bio-ethanol blend into gasoline is in usage nowadays (10). Model microorganisms have been used as the producing hosts, including *E.coli* (11-13) and yeasts (14-17) for their well-known genetic information, easy transformation and fermentation, as well as cyanobacteria (18-20) for they are *in vivo* aldehyde-producing.

Both modification of existing pathways and building synthetic routes via metabolic engineering are the common strategies. Steen *et al.* (12) engineered *E. coli* to produce structurally tailored fatty esters, alcohols and waxes. The hydrolysis of fatty acyl-ACP (acyl carrier protein), FA biosynthesis and FA carboxylate group reactivation were combined to divert FA to fatty acyl-ACP and fatty acyl-CoA, the important intermediates towards productions of potential biofuel. Atsumi *et al.* (13) engineered amino-acid biosynthetic pathway with high activities inside *E. coli*. The 2-keto acid intermediates were diverted into several alcohols from glucose. Engineering the existing metabolic pathways for alternation towards target compounds may prevent the possible metabolic imbalance and cytotoxicity which might be introduced by non-native pathway and accumulation of heterologous metabolites.

Tang *et al.* (14) used *S.cerevisiae* as the host and enhanced the cytosol acetyl-CoA production for FA biosynthesis. Blaceck *et al.* (16) reported a pentane producing pathway in *Y. lipolytica* by utilizing a soybean LOX enzyme to cleave linoleic acid.

This was the first microbial production of pentane and demonstrated the feasibility of short-chained *n*-alkane synthesis inside model microorganism hosts.

Kaiser *et al.* (18) discovered a cyanobacteria class-3 aldehyde dehydrogenase, whose overexpression resulted in 50~100-folds higher level of FAs than alkanes. The co-expression of one acyl-ACP reductase, one alcohol-dehydrogenase and one wax-ester synthase led to the yield of wax esters in the form of intracellular lipid bodies. Schirmer *et al.* (19) discovered one alkane biosynthesis pathway in cyanobacteria. The pathway consisted of one acyl-acyl carrier protein reductase and one ADC was found possible candidate to convert FA metabolism intermediates to alka(e)nes. Assumptions were proved by introducing the corresponding operons into *E. coli* and C₁₃ to C₁₇ alka(e)nes were produced.

Metabolic engineering now is a widely recognized methodology to produce natural and non-naturally existing product in microorganisms. It is foreseeable that along with the continuous development of systematic biology and extensive access to omics data, economic biofuel technologies with effective production hosts and efficient biomass conversions will be created which further lead to wider production and application of biofuels.

1.3 Fatty acids metabolism in *S. cerevisiae*

S. cerevisiae is genetically well-characterized. Apart from mature fermentation technique, its larger size leads to greater potential of production and better tolerance of exogenous genes. Clear understanding of metabolism further makes *S. cerevisiae* an ideal host in this research.

In *S. cerevisiae*, the cellular fatty acid (FA) pool is fed by: a) long-chain fatty acyl-

CoAs formed *de novo*; b) mobilization of endogenously stored FAs from lipids; and c) uptake of exogenous FAs (21). It has been reported that *S. cerevisiae* can survive with FAs as the sole carbon source (22). Although *S. cerevisiae* possessed a complete system to synthesize all required FAs, the ability of taking up FAs from the environment is indispensable. *S. cerevisiae* “prefer” importing FAs for usage due to the convenience of using existing materials instead of energy-consuming biosynthesis.

A lot of efforts have been devoted to identify the major route of FA degradation inside *S. cerevisiae*, and the genes and enzymes involved. It has been proved that FA β -oxidation procedure is completely confined to peroxisomes (23, 24). After absorption, FA activators would activate LCFAs into acyl-CoAs. The protein complex, Pxa1p-Pxa2p (Pat1p-Pat2p), embedded in the peroxisomal membrane, was revealed to function as a transporter and translocated the activated FAs into peroxisomes for degradation (Shown in Fig. 1.1, Page 7). The heterodimer Pxa1p-Pxa2p consist of Pxa1p and Pxa2p, both belong to the family of ABC transporters. Energy of ATP hydrolysis would be utilized for translocation across membranes. Previous research confirmed that *S. cerevisiae* $\Delta pxa1$, $\Delta pxa2$ and $\Delta pxa1\&2$ mutants can grow on medium contained LCFAs though growth was partially impaired as LCFAs cannot enter peroxisomes.

In this study, we used $\Delta pxa1$, $\Delta pxa2$ and $\Delta pxa1\&2$ mutants as the host of exogenous pathway. The exogenous genes would be introduced into the cytoplasm which was the platform of the exogenous pathway. The substrate of the pathway, LCFAs, also the sole carbon source in the culture would be absorbed and utilized by the strains.

Fig. 1.1 Scheme of β -oxidation in yeast peroxisomes (23). The protein complex, Pxa1p-Pxa2p (Pat1p-Pat2p) embedded in the peroxisomal membrane functions as a transporter, translocating activated FAs into peroxisomes for degradation. See descriptions on Page 6.

1.4 Proteomics

1.4.1 Introductions on proteomics

Proteome is the complete proteins expressed in one cell, tissue or organism.

Proteomics is the comprehensive analysis of proteome, including the large-scale

identification and localization, characterization of proteins including the structures and functions (25). It is complementary to genomics as it bridges gene sequences and proteins, the active agents in cells, and provided more direct understanding of post-translational modifications and interactions of proteins (26-28).

Proteomics is quite dynamic since protein levels may vary significantly at different time points due to the living conditions and status of the cell, whereas the genome of organisms is relatively constant.

Aiming to identify the proteins in samples of interest, proteomics makes it possible to characterize biological roles of proteins and activities of biological pathways. The common analytical process of proteomics is as follows: 1) crude protein sample collection; 2) sample resolution and specific labeling; 3) protein detection; 4) protein identification and quantification.

In order to fulfill the job of proteomics study, analyzing techniques for protein identification and quantification should be developed.

1.4.2 Basic introduction of LC-MS

The rapid advance of new experimental approaches has supported a rapid exploration phase of proteomics. Among these techniques, proteomics based on mass spectrometry (MS) has been a fundamental tool to interpret proteomics for “genome annotation” (29).

Numerous liquid chromatography (LC) techniques have been commercialized, such as size-exclusion chromatography, ion exchange chromatography, reversed phase chromatography and affinity chromatography. The high performance of LC

could have a wider range of applications from small-sized molecule metabolites to large-sized peptides and even to proteins with further improvement.

Mass spectrometry in combination with chromatographic separation techniques further improved the capabilities of each technique in terms of determination and identification, as did enhanced mass resolving and mass determining capabilities. Liquid chromatography-mass spectrometry (LC-MS) is one broadly useful technique for various applications due to the high sensitivity and selectivity. LC-MS is a technique in which analytes are first resolved chromatographically followed by introduction into the mass spectrometer. LC-MS applications are broad in scope, and include food, drug development, and proteomics.

One mass spectrometer consists of three components: an ion source, a mass analyzer, and a detector. The ion source first emits a beam of electrons with very high velocity and kinetic energy to collide and break for get molecules and generate charged fragments. Then the charged molecular fragments are transported to the mass analyzer and separated basing on their m/z ratios. Then, in the final step, the detector, records either the charge induced or the current produced when the ion passes by or hits its surface.

1.4.3 Proteomics basing on LC-MS

MS-based proteomics has proved itself to be an indispensable technique for genome interpretation. The capacity in comprehensive analysis abundant of proteins expressed in a cell is rapidly evolving supported by the progress of novel experimental techniques (29-31).

Four general types of mass analyzer are currently in use in proteomics research (29). These are ion trap, time-of-flight (TOF), quadrupole and Fourier transform ion cyclotron (FT-MS) analyzers.

The structures and performance of these analyzers are quite different give each their own strength and weakness. They can be stand alone or put together in tandem for taking advantage of the strengths of each (29). (Shown in Fig. 1.2, Page 12).

The principles of ionization and sample introduction process in electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) are shown in Fig. 1.2 (page 12), upper left and upper right.

Fig. 1.2a-1.2f shows 6 instrumental configurations with their typical source. Fig. 1.2a is the scheme of a reflector TOF apparatus (MALDI equipped). The ions are accelerated to very high speed and kinetic energy. The ions are separated basing on their velocities after fly through a flight tube. Then one reflector, which compensates for slight differences in kinetic energy, changes the route of the ions to impinge on the detector. MALDI-TOF mainly dedicated to rapid peptide mass fingerprint measurements. The mass range can be either well suited for tryptic peptides in reflectron mode, or up to few hundred kDa in linear mode (32).

Fig. 1.2b is the scheme of the TOF-TOF apparatus (MALDI equipped). One collision cell is incorporated between two TOF sections. Ions fly through the first TOF would be selected due to the m/z ratios. The collision cell fragments ions into fragments which are further separated by the following TOF part. MALDI-TOF-TOF is a MS/MS capable instrument with high-throughput (32).

Fig. 1.2c is the scheme of the quadrupole mass spectrometer/linear ion trap (ESI equipped). The time-varying electric field between 4 rods only permits ions with certain trajectory and particular m/z to travel through. After entering the instrument, the ions of a particular m/z value are first screened by the section Q1, followed by fragmentation in the collision cell (q2) and then the fragments are separated in Q3. ESI-triple quadrupole is a pretty old set-up nowadays, but still can work in specific measurement modes, like parent scanning or neutral loss. ESI-linear ion trap is a recent set-up which displays interesting features and will carry on evolving (32).

Fig. 1.2d is the scheme of a quadrupole-TOF apparatus, combined with a triple quadrupole instrument (front part) and one reflector TOF section to measure the ion masses (ESI equipped). ESI-Q-TOF, with high resolution, is commonly used as upstream in LC-MS/MS runs (32)

Fig. 1.2e is the scheme of 3-D ion trap (ESI equipped). It may capture the ions in the case of the linear ion trap, fragments ions of particular m/z ratio, and scans out the fragments to generate the tandem mass spectrum. It is capable of MS^n measurement as it allows fragmentation and mass analysis in the same space. The combination with LC is capable of high-rate acquisition, with high sensitivity and low resolution comparing to Q-TOF (32).

Fig. 1.2f is the scheme of the FT-MS apparatus combined with the linear ion trap for efficient and effective isolation, fragmentation and detection in the FT-MS sections (ESI equipped). FT-MS is an instrument of very high-resolution, high accuracy, high sensitivity and wide mass range; however, high price currently limited its usage (32).

Fig. 1.2 Mass spectrometers used in proteome research (29). The mechanisms of ionization by ESI and MALDI are shown in upper left and upper right. Fig. 1.2a-f shows the configurations of specific mass spectrometers. Fig. 1.2a: reflector TOF instrument. Fig. 1.2b: TOF-TOF instrument. Fig. 1.2c: triple quadrupole or linear ion trap. Fig. 1.2d: quadrupole-TOF. Fig. 1.2e: 3-D ion trap. Fig. 1.2f: FT-MS instrument. See descriptions on Page 10-11.

In this study, the LC-MS technique was adopted for proteomic analysis. Typical proteomics experiment procedure included the following five steps (29) (Shown in Fig. 1.3, Page 14).

In stage 1, the target proteins for analysis are extracted from cell lysates or tissues through biochemistry fractionation or affinity selection. The intact proteins may provide insufficient information for identification so they are digested by particular enzymes into linear peptides in stage 2. Trypsin is a common option for digestion, resulting in peptides with C-terminally protonated amino acids.

In stage 3, the peptides are separated by HPLC and introduced into an ion source to be nebulized into tiny, highly charged droplets. After evaporation, multiply protonated peptide was eluted into mass spectrometer and in stage 4, a mass spectra of the peptides mixture would be generated.

In stage 5, the computer provides data of the peptides which consist of the isolation of a given peptide ion, fragmentation by energetic collision with gas and record of the tandem or MS/MS spectra. The MS and MS/MS spectra are stored for matching against protein sequence databased. A list of identification of the peptides would be generated along with the credibility evaluation.

Fig. 1.3 Generic MS-based proteomics experiment (29). Stage 1: target proteins extraction. Stage 2: digest proteins to linear peptides. Stage 3: peptides chromatography, ionization and elution into MS. Stage 4, protonated peptides enters the MS and a mass spectrum would be generated. Stage 5: data analysis. See descriptions on Page 13.

1.4.4 Stable-isotope technique

To add a quantitative dimension to peptide LC-MS/MS experiments, the proven technique of stable-isotope labeling has been applied and commercialized (29).

This technique is based on the facts that the mixture of chemically identical analytes labeling with different stable-isotopes could be distinguished by MS detector, owing to their mass difference. The ratios of the signal isotope intensities could precisely indicate the abundance ratio of the two analytes (Fig. 1.4, Page 17).

The techniques to introduce stable-isotope tags to proteins includes metabolic stable-isotope labeling, isotope tagging by chemical reaction, and stable-isotope incorporation via enzyme reaction (Fig. 1.4, Page 17). In each case, the labelled proteins or peptides are combined together and the mixture was analyzed by MS for protein identification and relative abundance determination.

Fig. 1.4a shows the proteins labelled metabolically by culturing cells in medium which are rich or depleted in isotopes (N-isotope or C-isotope labelled). The advantage of metabolic incorporation is that the label is present in live cell. Separated labelled cells or tissues can be mixed before lysis. Moreover, the absence of side reactions (catalyzed by high specificity of enzyme) is another advantage (33).

Fig. 1.4b shows the proteins labelled at specific reactive sites with reagents containing stable isotopes. The prototypical study is the isotope-coded affinity tag reported in 1999 (34). Nowadays, quantitative tags with multiple reactive functional groups are available. The amino and carboxyl group-directed approaches can be applied to every peptide, in principle (33). Among them,

iTRAQ, the strategy targeting the amino group, recently gaining popularity, was used in this research (35). The mass spectra are relatively simple due to the properties of the labelling tag. Moreover, multiplexing strategy increased analytical throughput which allows 4 separately labelled proteins in one analysis.

Fig. 1.4c shows the stable-isotope tags were incorporated to proteins via enzyme reactions, i.e. ^{18}O from ^{18}O water during proteolysis (36, 37). Each peptide generated through the enzymatic reaction would carry the isotope tag, which is the foundation of analysis. However, one or both carboxyl oxygens may be exchanged which lead to variability into quantification, while only one ^{18}O atom incorporated was not sufficient for separation and quantification. As a result, the application range of this ^{18}O method was not wide (33).

The patterns of stable-isotope labelled mass differences generated by each method are shown in Fig. 1.4, Page 17.

Fig. 1.4 Scheme of stable-isotope protein labelling for quantitative proteomics

(29). Fig. 1.4a: Labeling proteins metabolically. Fig. 1.4b: Isotope tagging by chemical reaction. Fig. 1.4c: Stable-isotope incorporation via enzyme reaction. See descriptions on Page 15-16.

In this study, the isobaric tag for relative and absolute quantitation (iTRAQ) methodology has been adopted. In 2004, the multiplexed peptide quantitation methodology was reported (35). A set of 4 isobaric peptide reagents for peptide labeling for informative MS/MS spectra acquisition was used to qualify and quantify peptides (Fig. 1.5, Page 18).

Fig. 1.5 Principles of isobaric tagging by iTRAQ (35). Fig. 1.5A: the diagram of the structures of the multiplexed isobaric tags. Fig. 1.5B: the design of isotopes ^{13}C , ^{15}N , and ^{18}O . Fig. 1.5C: illustration of the isotopic tagging used to arrive at four isobaric combinations with four different reporter group masses.

Fig. 1.5A is the diagram of the structures of the multiplexed isobaric tags. The molecule contains one reporter group (based on N-methylpiperazine), one mass balance group (carbonyl) and one peptide-reactive group (NHS ester). The mass ranges of the reporter groups and balance groups were m/z 114–117 and m/z 31–28 respectively. The total mass of isobaric tag was kept constant as m/z 145 using different isotopic enrichment.

Fig. 1.5B shows that the design of label species containing isotopes ^{13}C , ^{15}N , and ^{18}O . It is noteworthy that the number and position of isotopic-enriched centers in the ring do not influence the chromatographic or MS behavior. When reacted with

a peptide, the tags will link to any peptide amine. These amide linkages fragment in a similar manner to backbone peptide bonds when subjected to collision-induced dissociation (CID). The fragmentation of the tag amide bond would lose the balance (carbonyl) moiety and retain the charge by the reporter group fragment.

Fig. 1.5C: Due to identical m/z , the mixture of 4 identical peptides labeled with one of the multiplex reagents would be resolved by LC as a single, unresolved MS precursor. Following CID, the 4 reporter group ions appear as distinct masses (114-117 Da). Furthermore, all other sequence-informative fragment ions remain isobaric. This remains the case even for those tryptic peptides that are labeled at both the N-terminus and lysine side chains, and those peptides containing internal lysine residues due to incomplete cleavage with trypsin. The relative concentration of the peptides is thus deduced from the relative intensities of the corresponding reporter ions.

Basing on these principles, the methodology of iTRAQ has been developed and industrialized (38-41). The isobaric nature of the tags enables simultaneous determination of multiple samples or samples containing internal standard. Furthermore, quantitative variability that may be introduced by individual experiment is removed by the mixture determination. Moreover, iTRAQ could be applied into a wide range of research, as the tagging chemistry is global and any peptide with a free amine can be labeled and measured. With above advantages, now iTRAQ is a widely used method in proteomics research. In this study, iTRAQ reagents were used to study the global protein profiles of multiple yeast strains.

1.5 Basic introduction of GC

GC is a commonly used analytical technique for gas samples or volatile compounds. Applications of GC include determination of the purity of a particular substance, separation and determination of relative amounts of different components of a mixture.

Generally, separation occurs when the sample is injected into the mobile phase, mostly an inert gas such as helium, which carries gaseous compound through stationary phase for analysis. The stationary phase is a column of metal tubing which contains an internal microscopic layer of liquid or polymer on an inert solid support. The gaseous compounds then elute out of the column at different time points due to different interactions in the column, i.e. the retention time, and captured by detectors.

One of the crucial components that determine the performance of GC is the column temperature. Here the column was contained in an oven, and temperature was precisely controlled electronically.

Several kinds of detectors are commonly used and were adopted in this study, including the flame ionization detector (FID) and MS detector.

FID works only for organic compounds. The detector includes electrodes placed adjacent to a flame fueled by hydrogen/air near the exit of the column (Fig. 1.6, Page 21). When organic compounds flow out of the column, they are pyrolyzed by the flame. The detector records the changes in current or voltage, which are translated as peaks in GC spectra.

For GC-MS, the MS spectrometry shares the mechanism and principles of LC-MS. The configuration of a typical GC-MS is shown in Fig. 1.7 (on Page 22).

GC-MS is a technique still in rapid development and improvement. Improvements include the following aspects: a) enhanced sample identification; b) extending the range of compounds amenable for GC-MS analysis; c) faster-speed GC-MS analysis; d) enhanced sensitivity; e) uniform, compound independent ion source response; f) better GC-MS compatibility with advanced technologies; g) higher GC-MS flexibility, ease of use and price; and h) better utilization of mass analyzer.

The Flame Ionisation Detector

Fig. 1.6 Scheme of the FID. FID works only for organic compounds. Two electrodes placed adjacent to a flame fueled by hydrogen/air near the exit of the column. Flow through of the organic compounds pyrolyzed by the flame would change the current or voltage between the electrodes, which were recorded as signals for translation into retention time. See descriptions on Page 20.

Fig. 1.7 The scheme of typical GC-MS. See descriptions on Page 21.

CHAPTER 2

OBJECTIVES

Unlike ethanol, hydrocarbons are highly compatible with existing energy infrastructure due to their chemical resemblance to petroleum-based fuels (42). Hydrocarbons are also energy-equivalent to petroleum-based fuels and therefore their usage results in no mileage penalty. Moreover, being immiscible in water eliminates the additional effort required, if any, for water separation and distillation step (42). This further makes medium- and long-chained hydrocarbons promising diesel substitutes. Fatty aldehydes derived from lipid biosynthesis were identified to be metabolically flexible precursors for a diversity of biofuels, including alkanes, free FAs, and so forth (4, 18).

Since hydrocarbons are potential biofuel to substitute fossil fuels, we therefore, in this study, devoted to the biosynthesis of medium-chained biofuel. We aim to find a “bio” approach to acquire renewable energies. With more devoted effort, exploration, improvement and development, it may support the energy consumption, ease the energy shortage, and relieve the reliance of development on fossil energy, which further lightens the stress laid on our environment.

This study described the construction of whole-cell based catalyst which was capable of producing medium-chained aldehyde as precursor for hydrocarbon biofuels, adopting the metabolic engineering techniques (Chapter 4). Moreover, the production of medium-chain alka(e)nes as hydrocarbon biofuels was also studied

preliminarily (chapter 5). One branch of the metabolism of oxylipin, the hydroperoxide pathway was found to be potential to fulfill the metabolic conversion (As shown in Fig. 2.1 below).

Fig. 2.1 Pathway of synthesis of medium-chained biofuel precursors.

S. cerevisiae, genetically well-characterized, with fast growth rates and convenient culture procedures, is an ideal microorganism for metabolic engineering. Moreover, clear understanding of lipid metabolic processes, including FA synthesis, elongation as well as FA degradation through β -oxidation, further makes *S. cerevisiae* an ideal host for the main objective of this study. We therefore used *S. cerevisiae* as the producing host, explored the biosynthesis of medium-chained aldehydes as hydrocarbon biofuel precursors through metabolic engineering techniques.

CHAPTER 3

RESEARCH DESIGN AND METHODS

3.1 Strains and media

E.coli strain Top10 was adopted for cloning and plasmid propagation. *E.coli* were cultured at 37°C with constant shaking at 250 rpm. Lysogeny broth contained 10 g/L bacto-tryptone (Fluka), 5 g/L yeast extract and 5 g/L NaCl (Sigma).

All the *S. cerevisiae* strains used in this study (shown in Table 3.1, Page 27-28) were cultured at 30°C with constant shaking at 250 rpm. YPD medium consisted 10 g/L yeast extract (Fluka), 20 g/L peptone (Bacto) and 20 g/L dextrose (Sigma). YNB-LEU selective media consisted of 6.7 g/L Yeast Nitrogen Base without Amino Acids (Sigma), 0.69 g/L DO Supplement-LEU (Clontech) and 20 g/L dextrose or galactose (Sigma). YNB-HIS selective media contained 6.7 g/L yeast nitrogen base without amino acids (Sigma), 0.69 g/L DO Supplement-HIS (Clontech) and 20 g/L dextrose or galactose (Sigma).

3.2 Double deletion strain construction

The pUG plasmid carrying gene disruption cassettes consisting of HIS5 heterologous marker genes with *loxP* sites was selected for gene disruption (43, 44). The *pxa1* and *pxa2* genes of *S. cerevisiae*, heterodimers of the peroxisomal membrane transporter, that translocates long-chained FA across the membrane, were targeted genes.

The sequences flanking the target genes in the genome were added to the 5' end of OL3' and OL3' sequences: 40 nucleotide stretches that are homologous to sequences upstream of the ATG start codon and down-stream of the stop codon of targeted gene, respectively. The detail primer sequences are shown in Table 3.1 (Page 27-28). The gene disruption procedure is shown in Fig. 3.1 (Page 29).

The *Δpxa1* strain was transformed with the *pxa2::his* while *Δpxa2* strain was transformed with the *pxa1::his* through PEG-LiAc method (45).

Fig. 3.1 (Page 29) also shows the principles of PCR confirmation for gene disruption. For the original strain, both combinations of primers A+B and primers C+D would synthesize an oligonucleotide sequence with a specific chain length, using PCR. For the disrupted strain, both combinations of primers A+ HisB, and primers HisC+D would synthesize an oligonucleotide sequence with a certain chain length. Otherwise, no oligonucleotide would be synthesized.

Primers for confirmation were designed and synthesized (Integrated DNA Technology), as shown in Table 3.2 (Page 30) and Table 3.3 (Page 31).

Table 3.1 Primers, plasmids and strains used into study

Name	Description	Reference
<i>E. coli</i> strain		
Top 10	F- <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) Φ 80 <i>lacZ</i> Δ M15 Δ <i>lacX74</i> <i>recA1</i> <i>araD139</i> Δ (<i>ara leu</i>) 7697 <i>galU galKrpsL</i> (StrR) <i>endA1 nupG</i>	Life Technology
Primers for gene disruption		
Δ <i>pxa2</i> -F	5'- ATAATAATAC AATTAAAAGT TACCGAAGAA AGATTTTATA CAGCTGAAGC TTCGTACGC-3'	This work
Δ <i>pxa2</i> -R	5'- CAATTTATAC ATGATTTGGA TCCTCCTTTG GCTATGTATG GCATAGGCCA CTAGTGGATC TG-3'	This work
Primers for cold fusion		
Cold fusion-F	5'- TATTTCTTGAATCAG TCAATATAGCAATGAGCAGT -3'	This work
Cold fusion-R	5'- GCCGAGCGGTCTAAG GAGCGACCTCATGCTATACC -3'	This work
Primers for restriction endonuclease		
F- <i>Bam</i> HI	5'-CGGGATCCAT GTTGCATAAC TTGTTGACA AGA-3'	This work
R- <i>Sal</i> I	5'-GCGTCGACAG TAGAATCCAA ACCCAACAAT GGA-3'	This work
Plasmids		
PUG27	<i>loxP-kanMX-loxP</i> disruption module plasmid	EUROSCARF
pESC- <i>leu</i>	With <i>GAL1</i> and <i>GAL10</i> yeast promoters in opposite orientation, <i>CYC1</i> and <i>ADH1</i> terminator respectively	Agilent
9LHP	9LOX-9HPL-pESC	This work
9LHPA	9LOX-9HPL-ADC-pESC	This work
pESC-ADC	ADC-pESC	This work

***S. cerevisiae* strains**

Wild type	<i>MATa; his3Δ 1; leu2Δ 0; met15Δ 0; ura3Δ 0</i>	EUROSCARF
<i>Δpxa1</i>	<i>BY4741; Mat a; his3D1; leu2D0; met15D0; ura3D0; YPL147w::kanMX4</i>	EUROSCARF
<i>Δpxa2</i>	<i>BY4741; Mat a; his3D1; leu2D0; met15D0; ura3D0; YKL188c::kanMX4</i>	EUROSCARF
<i>Δpxa1&2</i>	<i>BY4741; Mat a; his3D1; leu2D0; met15D0; ura3D0; YPL147w::kanMX4; YKL188c::his</i>	This work
WT-pESC	wild type carrying pESC	This work
<i>Δpxa1</i> -pESC	<i>Δpxa1</i> carrying pESC	This work
<i>Δpxa2</i> - pESC	<i>Δpxa2</i> carrying pESC	This work
<i>Δpxa1&2</i> - pESC	<i>Δpxa1&2</i> carrying pESC	This work
WT-9LHP	wild type carrying 9LHP	This work
<i>Δpxa1</i> -9LHP	<i>Δpxa1</i> carrying 9LHP	This work
<i>Δpxa2</i> -9LHP	<i>Δpxa2</i> carrying 9LHP	This work
<i>Δpxa1&2</i> -9LHP	<i>Δpxa1&2</i> carrying 9LHP	This work
WT-9LHPA	wild type carrying 9LHPA	This work
<i>Δpxa1</i> -9LHPA	<i>Δpxa1</i> carrying 9LHPA	This work
<i>Δpxa2</i> -9LHPA	<i>Δpxa2</i> carrying 9LHPA	This work
<i>Δpxa1&2</i> -9LHPA	<i>Δpxa1&2</i> carrying 9LHPA	This work

Fig. 3.1 Outline of the one-step gene disruption approach for generation of $\Delta pxa1\&2$ strain. See descriptions on Page 26.

Table 3.2 Primer designs to confirm gene disruption.

Name	Sequence
Primer His5B	5'- GGATGTATGG GCTAAATG -3'
Primer His5C	5'-CGTATGTGAA TGCTGGTC-3'
Primer A	5'- ACCGCTTGAA CCTGTGGA-3'
Primer B	5'- TTAGTTTCGC TGGTGCTG -3'
Primer C	5'- AAAGGTTGGG AAGATGAG -3'
Primer D	5'- AGAGTGCCCG TTAGGAAA-3'
Primer A1	5'-AGGAACAGAC GGAGTGG-3'
Primer B1	5'-ATGTATGAAG AACGCAGTAA-3'
Primer His5B1	5'-ACCTCAGTGG CAAATCC-3'
Primer A2	5'-GGTGTAGCCA CTGTATTG-3'
Primer B2:	5'-CTGTTCCCTG AATGTCG-3'
Primer His5B2:	5'-TTGTTTATGT TCGGATG-3'
Primer A3	5'-ACGGAGTGGT CGTTACAT-3'
Primer B3:	5'-CAAAGTTCCT TCACCCAT-3'
Primer His5B3:	5'-CAGTGGCAAA TCCTAACC-3'

Table 3.3 The combinations of primers used in PCR and resulting product sizes.

Primer combinations	Tm(°C)	Product size
A+B	42.0	499
A+ His B	43.3	771
C+D	41.9	259
HisC+D	42.2	468
A1+B1	43.8	871
A1+ His B1	47.0	1205
A2+B2	44.8	2687
A2+ His B2	39.0	1456
A3+B3	44.8	1626
A3+ His B3	45.7	1193
A1+D-ORF	41.9	3475
A1+D-His5	41.9	2367
A2+D-ORF	40.2	3655
A2+D-His5	40.2	2547
A3+D-ORF	41.0	3467
A3+D-His5	41.0	2359

Transformed deletion stains $\Delta pxa1$ & 2 strain were selected by histidine prototrophy by growing on synthetic complete minimal medium deficient in histidine (YNB-HIS) and yeast colony PCR was carried out for confirmation.

3.3 Cloning target genes

3.3.1 LOX and HPL cloning

All the oligonucleotide primers in Table 3.1(Page 27-28) were synthesized by Integrated DNA Technologies. All restriction enzymes, including *Bam*HI^{HF}, *Not*I^{HF}, *Nhe*I^{HF}, *Pac*I, *Sac*I^{HF}, *Sal*I^{HF}, *Xho*I were purchased from New England Biolabs. Ligation reactions were performed using T4 ligase (Fermentas). PCR reactions were carried out with HotStarTaq *Plus* Master Mix Kit (Qiagen) according to standard protocols, using a thermal cycler (MultiGeneTM OptiMax, Labnet). Gel extractions were carried out with QIAquick Gel Extraction Kit (Qiagen). *E.coli* minipreps were performed with QIAprep Spin Miniprep Kit (Qiagen).

Condon optimized genes 9LOX (V5 tag) and 9HPL (HA tag) were generated by Geneart (acc. No. KC920894/ KC920895). pESC-*leu* from Agilent was chosen as the cloning vector, which contains the *GAL*I and *GAL*10 yeast promoters in opposing orientations, capable of introducing two genes into one yeast host strain under the control of a repressible promoter. pES-*leu* contains a selective gene marker and resistive to ampicillin. Ampicillin sodium salt (Sigma) was added to LB for selection.

Primers F-*Bam*HI and R-*Sal*I in Table 3.1 (on Page 27-28) were used to introduce *Bam*HI and *Sal*I to 9LOX. Flanked by 5' *Bam*HI restriction site and 3' *Sal*I site, 9LOX gene was then inserted into pESC plasmid to obtain pESC-9LOX recombinant plasmid. *Sac*I and *Not*I restriction endonucleases were adopted to double digest 9HPL gene from default plasmid pMK-RQ. Then the DNA fragment,

Flanked by 5' *SacI* restriction site and 3' *NotI* site, was inserted into pESC-9LOX recombinant plasmid to obtain recombinant plasmid pESC-9LOX-9HPL (9LHP) (Fig. 3.2, Page 33). Four different double digestions were designed to confirm the recombinant vector 9LHP (11820 bp): *SacI* (3320) and *SalI* (8160); *PacI* (3312) and *SalI* (8160); *SacI* (3320) and *XhoI* (8196); *PacI* (3312) and *NheI* (8218). Gene sequencing was carried out to check the sequence of the recombinant plasmid (1st BASE Ltd).

Fig. 3.2 Scheme of recombinant plasmid 9LOX-9HPL-pESC (9LHP), 11820 kb. See descriptions on Page 32-33.

3.3.2 ADC cloning

Restriction enzymes, including *Clal*^{HF}, *NheI*^{HF}, *KasI*, *SacI*^{HF}, *SpeI*^{HF}, *PacI*^{HF}, were purchased from New England Biolabs. PCR reactions were carried out with HotStarTaq *Plus* Master Mix Kit (Qiagen) according to standard protocols.

The codon optimized ADC gene (*Synechococcus elongates*) with V5 tag was generated by GeneScript. Flanked by 5' *SpeI* restriction enzyme site and 3' *PacI* site, ADC was first inserted into empty vector pESC-*leu*. The primers for cold fusion PCR used to obtain the PCR insert, with GAL10 promoter, ADC gene and ADH1 terminator (1331 bp) were designed following the standard protocol of cold fusion PCR kit (System Biosciences), with 15 bp of vector sequence and 20 bp gene-specific sequence (Table 3.1, Page 27-28).

Exogenous gene ADC was fused into the recombinant plasmid 9LHP. *KasI* (2204) on recombinant plasmid were chosen to linearize 9LHP (shown in Fig. 3.2, Page 33). Cold fusion PCR was carried out following the standard protocol (45). We prepared the 10uL cloning reaction system of 1μL of 100 ng/μL linearized destination vector, 1uL of 200 μg/μL PCR insert, ddH₂O 7 μL and 2 μL 5× master mix. 50 μL cold fusion competent cell was added to the cloning mixture. After incubation on ice for 20 min, 50 s of heat shock at 42°C, 2 min of incubation on ice, 1h of incubation at 37°C with 250 μL LB broth added, the culture was spread on LB plate containing Ampicillin and cultured at 37°C overnight. The size of the recombinant plasmid 9LHPA was 13117 bp (shown in Fig. 3.3, Page 35). Two different combinations of double digestion were designed to confirm the recombinant vector 9LHPA (13117 bp): *SacI* (4627) and *NheI* (9525), *Clal* (1604) and *NheI* (9525).

Fig. 3.3 Scheme of recombinant plasmid 9LOX-9HPL-ADC-pESC (9LHPA), 13117 bp. See descriptions on Page 33-34.

3.4 Construction of functional *S. cerevisiae* strains

S. cerevisiae wild type and mutants were cultured in YPD. Frozen competent cells with high efficiency were produced following the protocol (45) as below:

Culture the yeast cells in a shaking incubator until the cell titer is at least 2×10^7 cells /mL. Harvest the cells and wash with sterile water twice; re-suspend the cells in 0.01 volumes of filter sterile frozen competent cell solution (5% v/v glycerol, 10% v/v DMSO); dispense into 1.5 mL microcentrifuge tubes; store in -80°C freezer overnight.

Transformations were carried out using LiAc/single strand carrier DNA/ PEG method (45) as follows:

Thaw competent cells and centrifuge for 2 min and remove the supernatant. Total volume of 360 mL of frozen competent cell transformation mix was prepared: 260 mL PEG 3350 (50% w/v), 36 mL LiAc (1 M), 50 mL single-stranded carrier DNA (2.0 mg/mL), and 14 mL plasmid DNA plus sterile water. Vortex mix to re-suspend the cell pellet; incubate in a 42°C water bath for 45 min; centrifuge and remove the supernatant; add 1 mL of YPD liquid medium into the transformation tube; vortex mix to re-suspend the cell pellet; spread on selective plates and culture.

The recombinant plasmid 9LHP was transformed into *S. cerevisiae* wild type, $\Delta pxa1$, $\Delta pxa2$ and $\Delta pxa1\&2$, to obtain WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1\&2$ -9LHP functional strains. Corresponding controls, WT-pESC, $\Delta pxa1$ -pESC, $\Delta pxa2$ -pESC and $\Delta pxa1\&2$ -pESC were constructed by transforming empty pESC plasmid into the four strains (Table 3.1, Page 27-28). Colony PCR was carried out to confirm the transformation.

The recombinant plasmid 9LHPA was transformed into *S. cerevisiae* wild type, $\Delta pxa1$, $\Delta pxa2$ and $\Delta pxa1\&2$, to obtain WT-9LHPA, $\Delta pxa1$ -9LHPA, $\Delta pxa2$ -9LHPA and $\Delta pxa1\&2$ -9LHPA functional strains (Table 3.1, Page 27-28). All the functional strains and control strains were confirmed by colony PCR and cultured on YPD and YNB-LEU selective minimal media.

The primers used in colony PCR are shown in Table 3.4 below.

Table 3.4 Primers for colony PCR

Name	Primer sequence	Product size (bp)
For pESC	F: 5'-TGTTCACTATCCCAAGCG-3'	1323
	R: 5'-CCCTATCTCGGTCTATTCT-3'	
For HPL	F: 5'-TTGCCATTGAAACCTATT-3'	730
	R: 5'-ATTCGTAGAAAGCATCGT-3'	
For LOX	F: 5'-CTTGCCATCTGAAACTCC-3'	1145
	R: 5'-TTGGATGAACGACGGAC-3'	
For ADC	F: 5'-AAACTTCCAAACTGCTGC-3'	460
	R: 5'-TGTACTGTCCAAACCCAAC-3'	

3.5 Growth curve test

Single colonies of the constructed functional strains and controls were selected into 50 mL tubes including 5 mL YPD with appropriate antibiotics. After growing overnight, 1×10^8 cells were injected into 250 mL flasks including 50 mL YPD medium. 1 mL was collected or diluted to the suitable OD value for sampling every 4 hours and the OD₆₀₀ were measured to determine the cell density. All of the cultures were maintained at 30°C, 250 rpm.

3.6 Protein extraction and Western-blot analysis

Western blot is a common technique for detection of specific proteins. As the promoter in the constructed plasmid was dextrose depressing and galactose

inducing, galactose, as carbon source and promoter inducer, was added into YNB cultures with containing ampicillin as selective marker. After two days culture, yeast cells were harvested and lysed using the glass beads method (14). Equal volumes of acid-washed glass beads (425-600 μm ; Sigma) and yeast lysis buffer (50 mM HEPES, 5% glycerol, 1 mM DTT, 1 mM PMSF, 1 mM EDTA) were added to extract crude protein from yeast cells.

After vortexing and 10 min centrifugation at 13,000 rpm, clear supernatants were collected. Protein concentrations were determined using a 2D-quant kit (GE Healthcare) following the standard protocol as follows.

Prepare standard curve of 0, 10 μg , 20 μg , 30 μg , 40 μg , 50 μg using the 2 mg/ml Bovine serum albumin (BSA) standard solution. Add 500 μL precipitant to each sample and standard curve tubes and vortex briefly to mix. Incubate the tubes 2-3 min at room temperature. Then add 500 μL co-precipitant to each tube and vortex briefly to mix. Centrifuge the tubes at a minimum of 13,000 rpm for 5 min. Decant the supernatants completely till no visible liquid remaining in the tubes. Add 100 μL of copper solution and 400 μL of distilled water to each tube and vortex briefly to dissolve the precipitate. Then add 1 mL of working color reagent (100 parts of color reagent A with 1 part of color reagent B) to each tube. Mix by inversion and incubate at room temperature for 15-20 min. Read the absorbance of each sample and standard at 480 nm using water as the reference. Calculate the concentrate of protein sample again the standard curve.

3.6.1 Gel electrophoresis

Mini-PROTEIN Tetra cells were utilized to perform sodium dodecyl sulphate-poly

acrylamide gel electrophoresis (SDS-PAGE). All the accessories for the SDS-PAGE including casting stands, casting frames, gel releasers, glass plates were purchased from Bio-Rad.

The SDS-PAGE gel in a single electrophoresis run can be divided into stacking gel and separating gel. In this study, 8% separating gel was adopted. The compositions of the stacking gel and separating gel are shown in Table 3.5 below and Table 3.6 (Page 40). The stocking solutions were prepared as follows:

Tris/SDS pH 6.8 (4×): 6.05 g/40 mL water adjust to pH 6.8 by HCL, add water to 100 mL total volume, add 0.4 g SDS.

Tris/SDS pH 8.8 (4×): 91 g/300 mL water adjust to pH 8.8 by HCL, add water to 500 mL total volume, add 2 g SDS.

Table 3.5 SDS-PAGE protocol of separating gel

8% separating gel	Volume (mL)
30 % acrymide	4
4× Tris pH 8.8	3.75
ddH ₂ O	7.25
10 % ammonium persulfate (APS)	0.1
N,N,N',N'-Tetramethylethylenediamine (TEMED)	0.01

Table 3.6 SDS-PAGE protocol of stacking gel

Stacking gel	Volume (mL)
30 % acrymide	0.65
4× Tris pH 6.8	1.25
ddH ₂ O	3.05
APS	0.025
TEMED	0.005

After preparation of the separating gel, the solution was loaded in the glass plates fixed by casting frame and cast stand. Distilled water was added on the top of the solution to remove bubbles in the gel solution and ensure a load surface. After the polymerization was completed and the separating gel solution was solidified, distilled water was disposed of and stacking gel solution was added on the top of the gel. Then a 1.0 mm thick 10-well comb was carefully inserted between the glass plates without trapping air bubbles under the teeth. After 20 min at room temperature, the gel was ready for electrophoresis run.

The Mini-PROTEIN Tetra Cell was filled with SDS running buffer: 0.192 M glycine, 0.025 M Tris, 0.1 % SDS, pH 8.3. Then 100 µg of quantified protein samples were mixed with sample loading buffer (6×): 7 mL Tris/SDS pH 6.8, 3 mL glycerol, 1 g SDS, 0.93 g DTT, 1.2 mg bromophenol blue, add water to 10 mL. The mixture was boiled at 95°C for 5 min to denature the proteins.

Then the samples were loaded slowly into the wells to prevent any spill. 10 μ L Novex Sharp Pre-stained Protein Standard (Life technologies) was loaded as the protein marker with a molecular weight range from 3.5 kDa to 260 kDa.

Constant voltage was applied to the SDS-PAGE. The constant voltage at 80 V was for the loaded sample to travel through stacking gel. Then constant voltage at 150 V was to resolve sample proteins in the separating gel. The run was terminated when the loading dye almost reached the bottom of the gel.

3.6.2 Membrane transfer

After electrophoresis, the stacking gel was cut and disposed and the separating gel was soaked in protein transfer buffer which contains 1.82 g/L Tris and 8.65 g/L glycine.

Then the gel was electro-transferred to PVDF membrane (Amersham Hybond-P, GE healthcare) at a constant voltage of 21 V for 1 h (Semi-dry transfer cell, Bio-rad). The PVDF membrane was soaked in methanol for 10 s and for equilibrated in the protein transfer buffer for 10 min. Filter paper was cut to a suitable size and soaked in the transfer buffer for 10 min. The layout of the transfer system was filter paper/SDS-PAGE gel/PVDF membrane/filter paper from top to bottom. The transfer was performed at constant current of 21 mA for 70 min. The transfer efficiency could be estimated by the appearance of the pre-stained marker on the PVDF membrane.

3.6.3 Blocking

5% non-fat milk TBS buffer (25 mM Tris-HCl, pH 7.4, 150 mM NaCl) solution was added as blocking buffer to block the PVDF membrane to prevent non-specific background bindings.

3.6.4 Detection

The exogenous gene 9LOX and 9HPL were designed with V5 and HA tags, respectively. In order to identify the corresponding proteins, the anti-V5-HRP and mouse anti-HA-HRP antibodies from Invitrogen were used, respectively. Goat anti-mouse IgG (H+L) secondary antibody HRP conjugate (Thermo scientific) was used.

After gently rotating for 1 h, the PVDF membrane was incubated in 3% non-fat milk TBS buffer solution with 1.0 µg/mL corresponding primary antibody for immune-detection and rotation at 4°C overnight was conducted. Then the membrane was washed 3×10 min in TBST washing buffer (25 mM Tris-HCl, pH 7.4, 150 mM NaCl, 0.05% Tween-20) to remove the superfluous primary antibody. Afterwards, 3% non-fat milk TBS buffer with corresponding HRP-conjugated secondary antibody was incubated with the PVDF membrane.

After incubation at room temperature for 1 h, the membrane was washed three times with TBST washing buffer and signals were visualized using the SuperSignal™ West Pico chemiluminescence (ECL) substrate (Thermal Scientific). The G:BOX chemi fluorescent & chemiluminescent imaging system (SYNGENE) was adopted to capture the signals.

We also used western blot to double-check the results of proteomics study. PGK1p and GAPDHp were selected as internal controls. The anti-PGK 1 antibody and anti-GAPDH antibody loading control from Abcam were adopted. Goat anti-rabbit IgG (H+L) secondary antibody HRP conjugate (Thermo scientific) was used.

3.7 Biotransformation and identification

3.7.1 GC-FID

Functional yeast strains were cultured in baffled shake flasks till the cell density reaches $OD_{600}=1$. 20 mL of the culture was collected and prepared as resting cells by two times' wash with 100 mM potassium phosphate buffer (pH=6.5). Then resting cells were transferred to a 250 mL GL-45 Erlenmeyer flask (Chemglass Life Sciences). 20 mL potassium phosphate buffer and 100 μ L substrate solution (5% linolenic acid water solution with 0.2% tween-80) was added to start the biotransformation and sealed with GL-45 open top cap and parafilms (Chemglass Life Sciences). Flasks were incubated at 30°C on an orbital shaker (250 rpm) for 3 days.

1 mL SampleLock syringe (Hamilton) was adopted to inject the headspaces of the 20 mL cultures into GC system to determine aldehyde production. Analysis was done using an Agilent 6890N GC-FID system (Agilent) equipped with Agilent J&W DB-WAX column (30 m $\times$ 0.25 mm $\times$ 0.25 μ m, Agilent). GC settings were as follows: carrier gas: helium; column flow 2.0 mL/min; splitless; inject temperature 230°C. The analyzing temperature program used was as follows: 50-230°C in 18 min; 230°C for 2 min (46). Identification and quantification of products was carried out by comparing to authentic standards with benzaldehyde as the internal control.

3.7.2 GC-MS

The GC-MS analysis was performed with a Shimadzu QP2010 Plus system (Shimadzu, Kyoto, Japan), equipped with a Agilent J&W DB-WAX column (30 m $\times$ 0.25 mm $\times$ 0.25 μ m, Agilent). 1 mL of culture headspace was injected into the system. Helium was used as carrier gas and the flow rate was 2 mL/min. The solvent cut off time was 3 min. The injection temperature was kept at 250°C. The oven temperature program was as follows: 50-230°C in 18 min; 230°C for 2 min. The ion source temperature was 200°C. The detection was performed in an electron impact ionization mode at 70 eV. The mass spectrum was recorded from 35 to 350 m/z, with a 0.5 s scan time. The Shimadzu GC-MS solution software was applied to acquire the chromatogram and identify mass spectra. Prior to the peak area integration, the noise reduction and baseline correction were operated. The quantification of FAs was achieved by integrated peaks through comparing with the added internal standards and the components were identified from the standard and the metabolite mass spectra database.

3.8 Proteomics

The whole procedure of cell lysis and protein extraction was carried out on ice to prevent protein denaturation. Same amount of yeast cells ($OD_{600}=20$) were pelleted at 13,000 rpm, 4°C, for 5 min. The cell pellets were washed twice with distilled water and re-suspended in 300 μ L of yeast lysis buffer which consisted of: 8 M Urea, 50 mM DTT, 50 mM Tris-Cl (pH7.6), 100 mM NaCl, 0.1% Triton X-100, 1 mM EDTA and 1 mM PMSF. An equal volume of acid-washed glass beads was added and the mixture was treated a bead mill using 4 cycles of 30s of vortexing at

4.0 m/s, with 30s of cooling on ice. Lysate was centrifuged at 10,000 rpm for 10 min at 4°C. The supernatant was collected and stored at -80°C. Protein concentrations were determined by 2-D quant kit (47).

A total of 100 µg proteins from functional strains WT-9LHP, *Δpxa1*-9LHP, *Δpxa2*-9LHP and *Δpxa1&2*-9LHP were collected and labeled using an iTRAQ Reagent Multi-Plex Kit (AB Sciex). Each sample was added to 20 µL of dissolution buffer and 1 µL denaturant and vortexed to mix. Then 2 µL of reducing reagent was added. After incubation at 60°C for 1 h, 1 µL cysteine-blocking reagent was added, mix by vortexing and then incubate 10 min at room temperature. 20 µL of 0.25 µg/µL sequence grade modified trypsin (Promega, US) was added to each sample to digest the protein overnight at 37°C.

Then the iTRAQ reagents 114, 115, 116 and 117 were added to the samples respectively for labeling:

The WT-9LHP sample was labeled with iTRAQ tag 114; *Δpxa1*-9LHP with iTRAQ tag 115; *Δpxa2*-9LHP with iTRAQ tag 116; and the *Δpxa1&2*-9LHP with iTRAQ tag 117. The labeled samples were then combined and baked on a heater at 30°C for condensation to roughly 100 µL. The proteomics study provided us extent information of the overall protein status in yeast cells and evident to predict the biotransformation efficiencies of the functional strains.

Fig. 3.4 below shows the flowchart of iTRAQ-based quantitative proteomics experimental design.

Fig. 3.4 Flowchart of iTRAQ-based quantitative proteomics experimental design. See descriptions on Page 44-45.

3.9 Protein identification and data analysis

The labeled samples were analyzed using online 2D Nano-LC-MS/MS 1200 series nanoflow LC system (Agilent Technologies) interfaced with a 6500 Q-TOF mass-spectrometer with HPLC-Chip Cube (Agilent Technologies) for two dimensional analysis. The HPLC-Chip was a combination of Zorbax 300SB C₁₈ reversed-phase

column (75 μm ×50 mm, 3.5 μm) packed with Zorbax 300SB C₁₈ enrichment column (0.3×5 mm, 5 μm).

In the first dimension, 4 μL of sample was loaded onto a PolySulfoethyl A strong cation-exchange (SCX) column (0.32×50 mm, 5 μm). The retained peptides were then eluted by sequential injection of 8 μL salt plugs with series of ammonium formate solutions in 9 gradient concentrations of 20, 40, 60, 80, 100, 200, 500 and 1000 mM.

In the second dimension, the effluent from SCX column was trapped onto Zorbax 300SB C₁₈ enrichment column during the enrichment mode by buffer A (5% acetonitrile and 0.1% formic acid) with a flow rate of 4 $\mu\text{L}/\text{min}$. Then the HPLC-Chip was switched to analytical mode. The peptides previously trapped on the enrichment column were eluted for 60 min by buffer B (0.1% formic acid) and buffer C (a nanoflow gradient of 5% - 80% acetonitrile + 0.1% formic acid) at a flow rate of 300 nL/min. Subsequently, the effluent flowed through the analytical Zorbax 300SB C₁₈ reversed-phase column for separation (Fig. 3.5, Page 48).

The analysis was carried out using a 6500 Q-TOF mass spectrometer with a capillary voltage of 1950 V. In total, 10 runs were carried out to accomplish the analysis of one sample. For MS analysis, positive ionization mode was used. Survey scans were from m/z 300 to 2000 with an acquisition rate of 4 spectra/s (47).

Fig. 3.5 Workflow of on-line 2D nano-LC. (a) The peptides eluted through SCX column and trapped onto enrichment column. (b) HPLC-chip on analytical mode, the previously trapped peptides were eluted and analyzed by analytical column. See descriptions on Page 46-47.

3.10 Protein quantification and data analysis

Peptide/protein quantifications were performed with Spectrum Mill MS Proteomics Workbench (Agilent Technologies). MS/MS spectra were searched individually (for *S. cerevisiae* species) against the UniProt-Swiss-Prot database. Methyl-methane-thiosulfate-labeled cysteine and iTRAQ modification of free amine in the amino terminus and lysine were set as fixed modification. Protein relative quantification using iTRAQ was performed on the MS/MS scans. Protein

quantification data with two or more unique peptides identified with confidence > 99% and the p value < 0.05 were selected for further statistical analysis. Three independent batches were performed to increase statistical confidence of changes in protein expression. The overlapping isotopic contributions were used to correct the calculated peak area ratios and to estimate the relative abundances of a specific peptide.

Furthermore, Western blot was carried out to confirm the results of LC-MS. Representative proteins: phosphoglycerate kinase 1 and glyceraldehyde-3-phosphate dehydrogenase were selected as internal controls (47).

CHAPTER 4

METABOLICALLY ENGINEERED YEAST CELLS AND MEDIUM-CHAINED HYDROCARBON BIOFUEL PRECURSORS SYNTHESIS

4.1 Introduction

4.1.1 The lipoxygenase pathway

Plants have evolved a number of complex and precise signaling pathways to react to environmental stress to ensure their own normal growth. The metabolism of oxylipins, a group of compounds obtained from oxygenation of polyunsaturated fatty acids (PUFAs), is one major component of defense mechanisms employed in plants. At normal circumstances, the levels of oxylipins are quite low but a great and rapid increase occurs when plants undergo various environmental stresses (48).

PUFA oxygenation is catalyzed by LOX and the most common substrates are linoleic acid or linolenic acid. One peroxy is inserted position-specifically to the backbone of the FA to yield one hydroperoxide (HPOD). HPODs do not show significant biological activities but it has been proved that their downstream

metabolites are bioactive and function as signaling molecules. HPODs were subsequently further metabolized by several secondary pathways (49), including HPL, allene oxide synthase (AOS), divinyl ether synthase (DES), peroxygenase, and so forth. The lipoxygenase pathway is defined as the metabolic flux of the LOX-catalyzed oxygenation and the subsequent reactions of PUFAs.

Study of several HPLs, AOSs and DESs discovered that at the end of the pathway, different oxylipins including aldehydes, jasmonates, divinyl ethers and epoxy FAs, with specific biological functions respectively were produced. Among them, the HPL branch is the main focus of this work.

Plant pathways were classified into two groups according to different substrate specificities (50) as shown in Fig. 4.1 (Page 53). One group mainly works on the number 13 carbon atom, cleaves linoleic acid 13*S*-hydroperoxide (13-HPOD) or α -linolenic acid 13*S*-hydroperoxide (13-HPOT) into 12-oxo-(9*Z*)-dodecenoic acid and *n*-hexanal or (Z)-3-hexenal respectively. The other group works on the number 9 carbon atom, cleaves linoleic acid 9*S*-hydroperoxide (9-HPOD) or α -linolenic acid 9*S*-hydroperoxide (9-HPOT) into 9-oxononanoic acid and (Z)-3-nonenal or (Z, Z)-3, 6-nonadienal respectively.

Fig. 4.1 Diagram of short-chained aldehyde forming system in plants (51). See description on Page 52.

4.1.2 Lipoxygenase

LOXs are a family of non-heme-iron-containing enzymes existing prevalently in plants. They catalyze the region- and stereo-specific dioxygenation of PUFAs containing a (1Z, 4Z)-pentadiene structure (52). Basing on their positional specificities of oxygenation, plant LOXs were categories into 9LOXs and 13LOXs.

Previous work has studied the mechanisms of the positional specific production of these two hydroperoxy PUFAs regioisomers (52). Furthermore, other properties of LOXs including specificity, activity, kinetics and genetics from different sources have also been studied. The C-13 and C-9 positional-specific catalyzing activities mentioned above can be observed, for instance: hazelnut (53), English pea (54) which exhibited both activities; soybean(55), olive (56) revealed 13-LOX activity, while almond (57, 58) contains 9-LOX activity. This study mainly focuses on the LOXs from almond (*Prunus dulcis*) whose corresponding enzymes were reported to be 9- positional-specific.

Mita *et al.* (57) reported the active LOX existing in almond; the reaction product obtained with linoleic acid as substrate was almost entirely 9-HPOD. The LOX gene contains 862 amino acids and molecular mass of 98 kDa, comprising 9 exons separated by 8 introns. This was the first reported LOX which displayed high product specificity. Furthermore, Santino *et al.* (58) reported on the same gene, and expressed the cDNA of almond LOX in *E. coli* for biochemical characterization. The recombinant almond LOX showed similar biochemical features to native enzymes. Virtually 100% of 9-HPOD was produced either with linoleic acid or linolenic acids as substrate.

Thus, as the LOX from almond was established to be 9-positional specific, it agreed well with the objective of this study, and was selected for further analysis.

4.1.3 Hydroperoxide lyase

HPLs are enzymes which cleave the C-C bond adjacent to the hydroperoxy group in HPODs to yield ω -oxo acids and volatile aldehydes.

The mechanisms of catalysis and positional specificity of HPL have been proposed (50). HPL from tea leaves (59), green bell peppers (60), soybeans (61) are 13-HPOD specific while HPL from alfalfa (62), melon fruit (63) and cucumber (64) were reported to display both 9- and 13- hydroperoxide catalyzing activities.

One 9- position-specific HPL enzyme reported by Mita *et al.* (65) drew our attention. The author studied the properties of 9HPL and introduced it into *E. coli*. The volatile C₉ aldehydes were produced and expression up-regulation and the activity were characterized. The gene was proved to barely metabolize 13-HPODs, show strict 9-specificity, which was different from the ones from melon or cucumber reported before.

The co-expression of LOX and HPL in *S. cerevisiae* has been reported recently, mainly focused on converting linoleic acid into aldehydes and alcohols as aroma compounds (46), while the synthesis of fuels such as alka(e)nes using a similar approach has not been reported. As properties and quantities of aldehydes are closely related to aroma of plants, this pathway is widely studied and modified in agricultural and industrial applications, which provided extensive references.

4.2 Experiment procedure

In this chapter, we studied the hydrocarbon precursor biosynthesis with *S. cerevisiae* as the producing host. The exogenous genes were introduced into *S. cerevisiae* strains to construct functional strains and were further characterized.

We studied the proteomics of the functional strains. A WT-9LHP sample was labeled with iTRAQ tag 114; $\Delta pxa1$ -9LHP sample with iTRAQ tag 115; $\Delta pxa2$ -9LHP sample with iTRAQ tag 116; $\Delta pxa1\&2$ -9LHP sample with iTRAQ tag 117. The labeled samples were then combined and condensed to roughly 100 μ L.

The proteomics study provided us with extensive information of the overall protein status in yeast cells and evidence to predict the biotransformation efficiencies of the functional strains.

Biotransformation determination was carried out to detect the catalyzing activities of the whole-cell based catalysis.

4.3 Results

4.3.1 Construction of recombinant plasmid 9LHP

The 9LOX and 9HPL genes were fused into carrier plasmids and the size of the recombinant plasmid 9LHP was 11820 bp. Gene-sequencing results proved that no site mutation was present in the recombinant plasmid. Double digestion and electrophoresis were carried out to confirm this, as shown in Fig. 4.2 (Page 57). The gene sequencing results showed no site mutation in the final recombinant vector.

Fig. 4.2 Combined results of four different double digestions of recombinant vector 9LOX-9HPL-pESC (11820 bp) for confirmation. a) 5 k and 7 k with *SacI* (3320) and *SalI* (8160); b) 5 k and 7 k with *PacI* (3312) and *SalI* (8160); c) 5 k and 7 k with *SacI* (3320) and *XhoI* (8196); d) 5 k and 7 k with *PacI* (3312) and *NheI* (8218). (GeneRuler 1 kb Plus DNA Ladder, ready-to-use, Fermentas).

4.3.2 Construction of double deletion strain

We constructed double mutant $\Delta pxa1 \& 2$ using PUG27 plasmid carrying *lox-his5⁺-lox*. The primers were designed and PCR was repeated to confirm the gene disruption.

Fig. 4.3 (Page 58) shows the gel electrophoresis of gene disruption confirmation. For deletion strain, the ORF was disrupted and replaced by *his5⁺* tag. So PCR

using primers A+His5 and primers HisC+D should show bands of 771 bp and 468 bp respectively. PCR using primers A+B and primers C+D should show no band.

Fig. 4.3 The gel electrophoresis of gene disruption confirmation. Lane 1: A+His B, product size of 771 bp. Lane 2: A+B, band. Lane 3:C+D, no band . Lane 4:His C+D, product size of 468. See descriptions on Page 57.

Fig. 4.4 (Page 59) shows the gel electrophoresis to confirm the gene disruption. Several new primers were designed to repeat the confirmation of the gene disruption in case some unknown factors influenced the testing primers to shown positive results. For PCR using primers A+HisB1, A+HisB2 and A+HisB3, the sizes of the product oligonucleotide were 1205 bp, 1456 bp, and 1193 bp respectively.

Fig. 4.5 (on Page 59) shows the gel electrophoresis to repeat the confirmation of the gene disruption with different PCR primer combinations and designs. PCR was

carried out with different primer A and primer D. For PCR using primers A1+D, A2+D and A3+D, the sizes of the product oligonucleotides were 2367 bp, 2359 bp, and 1853 bp respectively.

Fig. 4.4 The gel electrophoresis of deletion strain. Lane 1: A1+His B1, product size of 1205 bp. Lane 2: A1+B1, no band. Lane 3:A2+His B2, product size of 1456 bp. Lane 4, A2+B2, no band. Lane 5:A3+His B3, product size of 1193. Lane 6, A3+B3, no band. See descriptions on Page 58.

Fig. 4.5 The gel electrophoresis f deletion strain. Lane 1/2: A1+D, product size of 2367 bp. Lane 3/4: blank. Lane 5/6: A2+D, product size of 2359. Lane 7:A3+D, product size of 1863. See descriptions on Page 58.

Thus, the double deletion strain $\Delta pxa1\&2$ was constructed. The strain was cultured with both YPD and YNB-HIS selective minimal plates. The strain grew well on both media (data not shown).

4.3.3 Construction of functional strains

The recombinant plasmid 9LHP or empty vector pESC were transformed into *S. cerevisiae* wild type, single mutant $\Delta pxa1$ and $\Delta pxa2$, and double mutant $\Delta pxa1\&2$ to obtain functional strains WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP, $\Delta pxa1\&2$ -9LHP and controls WT-pESC, $\Delta pxa1$ -pESC, $\Delta pxa2$ -pESC, $\Delta pxa1\&2$ -pESC, respectively.

Colony PCR was carried out to confirm the existence of the recombinant plasmid using the primers, shown in Table 3.4 (Page 37).

Fig. 4.6 (on Page 61) shows the colony PCR results against LOX gene and HPL gene. The size of the bands was 1145 bp and 730 bp respectively. This proved that the transformation of the recombinant plasmid 9LHP into the strains was successful.

Fig. 4.7 (on Page 61) shows the colony PCR results against pESC empty vector. The size of the band was 1323 bp. This proved that the transformation of the empty vector into the strains to construct control strains was successful.

Fig. 4.6 Colony PCR results of the functional strains. Left: LOX, size=1145 bp; Right: HPL, size= 730 bp. Lane 1: WT-9LHP. Lane 2: $\Delta pxa1$ -9LHP. Lane 3: $\Delta pxa1$ -9LHP. Lane 4: $\Delta pxa1 \& 2$ -9LHP. See descriptions on Page 60.

Fig. 4.7 Colony PCR results of the control strains against pESC, size= 1323 bp. Lane 1: WT-pESC. Lane 2: $\Delta pxa1$ - pESC. Lane 3: $\Delta pxa1$ - pESC. Lane 4: $\Delta pxa1 \& 2$ - pESC. See descriptions on Page 60.

4.3.4 Growth curve test

The OD values were measured every time of culture. For the time points 0 h, 4 h and 8 h, the OD values were measured without dilution. For the time points 12 h, 16 h, 20 h, 24 h, 28 h, and 32 h, the OD values were measure after 10 times dilution.

The growth curves of functional strains and controls were tested in YPD medium (Fig.4.8 below, full data shown in Table S1, Page 111). The functional strains grow slightly slower than controls. There was no significant growth variance inside the functional strain group and the control group.

Fig. 4.8 Growth curve of *S. cerevisiae* strains WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP, $\Delta pxa1 \& 2$ -9LHP and WT-pESC, $\Delta pxa1$ -pESC, $\Delta pxa2$ -pESC, $\Delta pxa1 \& 2$ -pESC as controls.

4.3.5 Western blot

To explore whether the 9LOX and 9HPL genes were effectively expressed, crude protein samples were extracted from yeast wild type and mutants for western blot according to the method introduced in Chapter 3 and immunoblotting analysis was also carried out for the functional strains: WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1\&2$ -9LHP with WT-pESC, $\Delta pxa1$ -pESC, $\Delta pxa2$ -pESC and $\Delta pxa1\&2$ -pESC as controls (Novex protein ladder, Invitrogen). Results shown in Fig. 4.9 indicated that both exogenous genes 9LOX and 9HPL can be successfully expressed in functional strains.

Fig. 4.9 Western blot of target genes in *S.cerevisiae* control strains and functional strains: a) 9LOX, 99kDa; b) 9HPL, 55kDa. Lane 1: WT-pESC. Lane 2: $\Delta pxa1$ -pESC. Lane 3: $\Delta pxa2$ -pESC. Lane 4: $\Delta pxa1\&2$ -pESC. Lane 5: WT-9LHP. Lane 6: $\Delta pxa1$ -9LHP. Lane 7: $\Delta pxa2$ -9LHP. Lane 8: $\Delta pxa1\&2$ -9LHP.

4.3.6 Proteomics

The proteomic profiling of *S.cerevisiae* strains was carried out by on-line 2D LC-MS/MS system. The Spectrum Mill system was used for peptide identification.

We studied proteomics of the functional strains WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1\&2$ -9LHP. Based on the analysis conditions, more than 200 proteins were detected. Raw data was shown in Fig. S1 (on Page 115).

The peptide fragmentation spectra of glucose-6-phosphate isomerase was selected as an example as shown in Fig. 4.10. Peptide summary of glucose-6-phosphate isomerase is shown in Fig. S3 (Page 117). MS digest results of glucose-6-phosphate isomerase are shown in Fig. S4-S8 (Page 118-122).

Fig. 4.10 Representative peptide fragmentation spectrum of glucose-6-phosphate isomerase: (R)AVYHVALR(N) in WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1\&2$ -9LHP combined sample. More details in Fig. S2, Page 121.

After analysis, 31 proteins displayed different levels among the four functional strains (Table 4.1, Page 67-69). Based on their catalyzing functions and pathway involved, we classified them into 8 categories:

a) galactose metabolism: galactokinase (EC 2.7.1.6) and galactose-1-phosphate uridylyltransferase (EC 2.7.7.12).

b) glycolysis: hexokinase-1 (EC 2.7.1.1), glucose-6-phosphate isomerase (EC 5.3.1.9), phosphofructokinase (EC 2.7.1.11), fructose-bisphosphate aldolase (EC 4.1.2.13), triosephosphate isomerase (EC 5.3.1.1), glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12), phosphoglycerate kinase (EC 2.7.2.3), phosphoglycerate mutase 1 (EC 5.4.2.11), enolase (EC 4.2.1.11) and pyruvate kinase (EC 2.7.1.40).

c) TCA cycle: citrate synthase, mitochondrial (EC 2.3.3.16) and aconitase, mitochondrial (EC 4.2.1.3).

d) ATP synthesis: ATP synthase subunit alpha, mitochondrial and ATP synthase subunit beta, mitochondrial (EC 3.6.3.14).

e) amino-acid metabolism: 3-isopropylmalate dehydratase (EC 4.2.1.33), 3-isopropylmalate dehydrogenase (EC 1.1.1.85), 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (EC 2.1.1.14) and pyruvate decarboxylase isozyme (EC 4.1.1.1).

f) protein biosynthesis: ATP-dependent RNA helicase eIF4A (EC 3.6.4.13), elongation factor 1-alpha, 60s ribosomal protein L4 and 60s ribosomal protein L19.

g) heat shock proteins: 12 kDa Heat shock protein, heat shock protein 26 and heat shock protein STI1.

h) others: mitochondrial outer membrane protein porin 1, S-adenosylmethionine synthetase 2 (EC 2.5.1.6), uncharacterized oxidoreductase YMR226C and superoxide dismutase [Cu-Zn] (EC 1.15.1.1).

Table 4.1 Relative changes in protein expression of *S. cerevisiae* WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1\&2$ -9LHP

Gene	Description	No. of peptides	Average of B/A	Average of C/A	Average of D/A
Galactose metabolism					
GAL1	Galactokinase	9	0.870 $\pm$ 0.340	0.587 $\pm$ 0.225	3.783 $\pm$ 0.215
GAL7	Galactose-1-phosphate uridylyltransferase	2	1.230 $\pm$ 0.005	1.387 $\pm$ 0.965	1.008 $\pm$ 0.021
Glycolysis					
HXK1	Hexokinase-1	2	0.723 $\pm$ 0.259	0.691 $\pm$ 0.104	3.489 $\pm$ 0.368
PGI1	Glucose-6-phosphate isomerase	3	1.032 $\pm$ 0.460	1.330 $\pm$ 0.180	9.891 $\pm$ 0.251
PFK2	Phosphofructokinase	3	1.236 $\pm$ 0.155	1.149 $\pm$ 0.100	1.426 $\pm$ 0.197
FBA1	Fructose biphosphate aldolase	8	1.094 $\pm$ 0.155	0.871 $\pm$ 0.235	2.304 $\pm$ 0.942
TPI1	Triosephosphate isomerase	6	1.123 $\pm$ 0.380	0.849 $\pm$ 0.070	1.038 $\pm$ 0.357
TDH	Glyceraldehyde 3-phosphate dehydrogenase	15	1.117 $\pm$ 0.305	0.925 $\pm$ 0.220	1.664 $\pm$ 0.541
PGK1	Phosphoglycerate kinase	19	1.066 $\pm$ 0.225	0.798 $\pm$ 0.120	1.286 $\pm$ 0.076
GPM1	Phosphoglycerate mutase 1	14	1.047 $\pm$ 0.210	0.750 $\pm$ 0.145	1.500 $\pm$ 0.457
ENO	Enolase	19	1.185 $\pm$ 0.160	0.912 $\pm$ 0.155	2.176 $\pm$ 0.478
PYK	Pyruvate kinase	6	1.186 $\pm$ 0.200	0.831 $\pm$ 0.195	1.831 $\pm$ 0.147
TCA cycle					

CIT1	Citrate synthase, mitochondrial	3	1.256±0.820	1.982±0.230	5.580±0.248
ACO1	Aconitase, mitochondrial	3	0.961±0.400	1.033±0.075	1.737±0.128
ATP synthesis					
ATP1	ATP synthase subunit alpha, mitochondrial	6	1.222±0.260	1.005±0.030	1.157±0.160
ATP2	ATP synthase subunit beta, mitochondrial	8	1.176±0.040	1.123±0.085	3.216±0.205
Amino-acid metabolism					
LEU1	3-isopropylmalate dehydratase	3	0.883±0.075	1.023±0.335	1.424±0.200
LEU2	3-isopropylmalate dehydrogenase	17	3.477±0.630	1.070±0.335	2.570±0.254
MET6	5-methyltetrahydropteroyltriglutamate-- homocysteine methyltransferase	10	1.050±0.125	0.881±0.430	2.018±0.121
PDC	Pyruvate decarboxylase isozyme	12	1.305±0.355	1.118±0.115	1.894±0.218
Protein biosynthesis					
TIF	ATP-dependent RNA helicase eIF4A	3	1.408±0.025	0.752±0.295	1.655±0.245
TEF1	Elongation factor 1-alpha	8	0.910±0.375	0.758±0.185	1.507±0.110
RPL4	60s ribosomal protein L4	9	1.245±0.255	0.778±0.035	1.418±0.068
RPL19	60s ribosomal protein L19	2	1.218±0.285	1.114±0.805	2.995±0.197
Heat shock proteins					
HSP 12	12 kDa Heat shock protein	2	2.199±0.640	0.882±0.135	2.308±0.219
HSP 26	Heat shock protein 26	3	2.281±0.675	1.823±0.360	2.453±0.195

STI1	Heat shock protein STI1	2	1.363±0.665	0.485±0.035	3.450±0.377
Unknown					
POR1	Mitochondrial outer membrane protein porin 1	4	1.033±0.395	0.808±0.445	2.785±0.066
SAM2	S-adenosylmethionine synthetase 2	2	1.271±0.300	0.624±0.085	2.125±0.151
YMR22 6C	Uncharacterized oxidoreductase YMR226C	2	1.856±0.375	1.260±0.110	3.051±0.265
SOD1	Superoxide dismutase [Cu-Zn]	2	6.360±0.420	3.942±1.400	7.910±0.330

The “Average of B/A” refers to the average ratio of protein expression level in $\Delta pxa1$ -9LHP strain over that in WT-9LHP strain. “Average of C/A” refers to the average ratio of protein expression level in $\Delta pxa2$ -9LHP strain over that in WT-9LHP strain. The “Average of D/A” refers to the average ratio of protein expression level in $\Delta pxa1 \& 2$ -9LHP strain over that in WT-9LHP strain. Average of protein expression levels in WT-9LHP strain was taken as 1 and the deviation was calculated from three independent LC-MS/MS analysis results.

A heat map of the proteomic results for better results clarification is shown in provided as Table S2 (on page 123).

In order to verify the results of LC-MS, Western blot was carried out for the representative proteins phosphoglycerate kinase 1 and glyceraldehyde-3-phosphate dehydrogenase as internal control (Shown in Fig. 4.11 below).

Fig. 4.11 Western blot results of phosphoglycerate kinase 1 (45 kDa) and glyceraldehyde-3-phosphate dehydrogenase (37 kDa) (A) Protein expression changes. (B) Quantification of protein level changes based on Western blot analysis.

The Western blot results are in accordance with LC-MS results.

It is noteworthy that the levels of the listed proteins in strains *Δpxa1*-9LHP and *Δpxa2*-9LHP were mostly equivalent to strain WT-9LHP, while in the *Δpxa1&2*-LHP strain, all the proteins listed were up-regulated to different extents (Table 4.1, Page 67-69).

As both the promoters *GAL1* and *GAL10* are inducible, galactose was added to the culture as the sole carbon source in order to induce the promoters on the recombinant plasmid pESC-*leu*. After absorption, galactose must be converted into glucose-6-phosphate (G6P) through Leloir pathway before it can enter glycolysis (shown in Fig. 4.12, Page 75).

Leloir is the metabolic pathway for galactose catabolism and contains 2 reversible reactions and 2 irreversible reactions (66). In yeast, the galactokinase is encoded by the *GAL1* gene and the galactose-1-phosphate uridylyltransferase is encoded by the *GAL7* gene (67). When galactose is the preferred carbon source, the transcription of *GALs* is induced rapidly by more than 1000-fold (68). Our proteomics results showed slight expression differences among WT-9LHP strain, *Δpxa1*-9LHP strain and *Δpxa2*-9LHP strain while up-regulation in *Δpxa1&2*-9LHP strain.

Galactose-converted G6P would then be degraded through glycolysis; this pathway contains 10 reactions to convert glucose to two C3 units (pyruvate) and release free energy in the process (shown in Fig. 4.13, Page 76). 3 of the reactions are candidates for flux-control points: those catalyzed by hexokinase,

phosphofructokinase and pyruvate kinase (69). In our case, galactose as the sole carbon source was converted into G6P so the hexokinase reaction is not required.

Table 4.1 (Page 67-69) showed that the levels of the 10 enzymes (including those catalyzing the rate-limiting reactions) in the $\Delta pxa1$ -9LHP and $\Delta pxa2$ -9LHP strains were similar to the WT-9LHP strain. The levels of all the ten proteins were up-regulated in $\Delta pxa1 \& 2$ -9LHP strain, which may suggest increased glycolysis.

Pyruvate is the end product of glycolysis. It can be used in aerobic respiration via the TCA cycle, a series of chemical reactions used to generate energy through the oxidation of acetate derived citrate (shown Fig. 4.14, Page 77). Pyruvate decarboxylated by pyruvate dehydrogenase was converted into acetyl-CoA, the starting point of TCA cycle. The rate-limiting steps are the reactions catalyzed by citrate synthase, isocitrate dehydrogenase and α -ketoglutarate dehydrogenase (69).

Table 4.1 (Page 67-69) showed that citrate synthase, catalyzing the rate-limiting step in TCA cycle, as well as aconitase, showed significantly higher levels in $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1 \& 2$ -9LHP strains than that in WT-9LHP strain, which may suggest up-regulated activity in the TCA cycle.

ATP synthase is the important enzyme that synthesizes ATP to provide energy for the cell. Located within mitochondria, ATP synthase consists of 2 regions: F_0 portion within the membrane and F_1 portion inside the matrix of the mitochondria. The ATP synthase subunit alpha and beta are the subunits of the F_1 -ATP synthase (70, 71). Our results showed that these two proteins were slight up-regulated in $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1 \& 2$ -9LHP strains relative to WT-9LHP.

The metabolic pathways mentioned above, galactose metabolism, glycolysis, TCA cycle and ATP synthesis, are steps in carbohydrate catabolism that release energy in the form of ATP. The enzymes involved in these pathways were notably up-regulated in the strain $\Delta pxa1$ & 2-9LHP which suggested that the metabolism and energy provision in this strain were more active and may lead to higher efficiency as a whole-cell catalyst.

With galactose inducing the promoters, exogenous genes 9LOX and 9LHP carried by high copy number vector pESC-*leu* were expressed in the four functional strains. LC-MS/MS results identified four enzymes involved in amino acid metabolism showed different expression levels: 3-isopropylmalate dehydratase, 3-isopropylmalate dehydrogenase, 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase and pyruvate decarboxylase isozyme.

3-isopropylmalate dehydratase and 3-isopropylmalate dehydrogenase are involved in leucine biosynthesis. 3-isopropylmalate dehydratase catalyzes the dehydration of 2-isopropylmalate to isopropylmaleate and then rehydration to 3-isopropylmalate (72). 3-isopropylmalate dehydrogenase catalyzes the oxidation of isopropylmalate to isopropyl isosuccinate, which undergoes non-enzymic decarboxylation to oxo-isocaproic acid, the precursor for transamination to leucine. The pESC-*leu* vector carrying the exogenous gene contains the selectable marker *LEU*. The strains were cultured in leucine-free media, and no available leucine from the medium. They have to self-generate leucine for all cellular processes including protein synthesis. 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase is involved in the biosynthesis of methionine (73, 74). Pyruvate decarboxylase isozyme is involved in the metabolism of valine (75). Our results

showed comparable expression levels among WT-9LHP, $\Delta pxa1$ -9LHP and $\Delta pxa2$ -9LHP strains, and while up-regulation in the $\Delta pxa1 \& 2$ -9LHP strain.

ATP-dependent RNA helicase eIF4F is involved in cap recognition and is required for mRNA binding to ribosomes (76, 77). Elongation factor 1-alpha, apart from a well-characterized role in translation elongation, its function in binding nascent polypeptide chains has also been reported (78). 60s ribosomal protein L4 and 60s ribosomal protein L19 play certain roles in cooperative ribosome protein functions (79). The above four proteins were also significantly up-regulated in the strain $\Delta pxa1 \& 2$ -9LHP which supported the expression of the exogenous genes well.

Furthermore, three heat shock proteins related to the stress response, 12 kDa heat shock protein, heat shock protein 26, heat shock protein STI1, were found to be up-regulated. The introduction of exogenous genes and the folding of the proteins may well stress the yeast cells, as up-regulation of these proteins in order to maintain balance of intracellular metabolism and the integrity of the whole cell.

In addition, the levels of mitochondrial outer membrane protein porin 1, S-adenosylmethionine synthetase 2, uncharacterized oxidoreductase YMR226C, superoxide dismutase [Cu-Zn] (80) were also found to differ among the four functional strains which remained to be studied.

The mechanism details of the above 31 proteins level differences however still need further investigation.

Fig. 4.12 Diagram of Leloir pathway (69). Galactokinase and galactose-1-phosphate uridylyltransferase catalyzing 2 irreversible reactions were identified in proteomics study. See descriptions on Page 71.

Fig. 4.13 Diagram of glycolysis (69). *hexokinase-1*, glucose-6-phosphate isomerase, *phosphofructokinase*, fructose-bisphosphate aldolase, triosephosphate isomerase, glyceraldehyde 3-phosphate dehydrogenase, phosphoglycerate kinase, phosphoglycerate mutase 1, enolase and *pyruvate kinase* were identified in proteomics (Enzymes catalyzing rate-limiting steps are in *italics*). See descriptions on Page 71-72.

Fig. 4.14 Diagram of TCA cycle (69). *Citrate synthase*, mitochondrial and *aconitase*, mitochondrial (Enzymes catalyzing rate-limiting steps are in *italics*). See descriptions on Page 72.

4.3.7 GC-FID detection

Functional strains and control strains were cultured in YNB-LEU selective media with galactose inducing the promoters of heterologous genes. As the produced 3(Z)-nonenal would be secreted and then, being insoluble in water phase and volatile, vaporize into the local environments, 1 mL of the headspace of the cultures was extracted and injected into GC-FID for qualification and quantification.

Preliminary results have shown that, when linoleic acid was added to cultures of the growing cells, no detectable targeting volatile compounds was produced (81). Thus non-growing but metabolically-active resting cells with higher specific catalyzing activities were adopted in this study.

Functional strains WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP, $\Delta pxa1\&2$ -LHP and corresponding control strains WT-pESC, $\Delta pxa1$ -pESC, $\Delta pxa2$ -pESC, $\Delta pxa1\&2$ -pESC were cultured, collected and prepared as resting cells for 3 days' biotransformation. Gas samples of the headspaces of the cultures were analyzed with GC-FID system.

Peaks at 8.82 min were identified as 3(Z)-nonenal by comparing the retention time of authentic standard (Fig. 4.15, Page 79). A series of different amounts of 3(Z)-nonenal were used to generate a standard curve for quantification. The characterizations of catalyzing activities of functional strains were repeated 5 times, and the control strains were repeated 3 times. Benzaldehyde was used as internal control for calibration. Fig. 4.16 (Page 80) shows the 3(Z)-nonenal production levels. All the control strains produced non-detectable levels of 3(Z)-nonenal.

Fig. 4.15 GC spectra: blue: 3(Z)-nonenal standard; red: $\Delta pxal\&2$ -9LHP; green: $\Delta pxal\&2$ -pESC. Retention time at 8.82 min was identified as 3(Z)-nonenal. See descriptions on Page 78.

Fig. 4.16 3(Z)-nonenal production levels by constructed strains. See descriptions on Page 78.

WT-9LHP, $\Delta pxa1$ -9LHP and $\Delta pxa2$ -9LHP strains displayed similar production capabilities, which were up to 0.57 ± 0.09 mg/L, 0.50 ± 0.07 mg/L and 0.48 ± 0.02 mg/L respectively, whereas $\Delta pxa1\&2$ produced twofold higher level of 3(Z)-nonenal, up to 1.21 ± 0.05 mg/L.

The catalyzing efficiencies of the functional strains were also characterized by the calculated carbon recovery rates. WT-9LHP, $\Delta pxa1$ -9LHP and $\Delta pxa2$ -9LHP have biotransformed 5.8%, 5.11% and 4.95% of linoleic acid into 3(Z)-nonenal respectively while $\Delta pxa1\&2$ -LHP showed the highest carbon recovery of up to 12.40% (Table 4.2 below).

Table 4.2 Production of functional strains and carbon recovery rates.

Functional strains	WT-9LHP	$\Delta pxa1$ -9LHP	$\Delta pxa2$ -9LHP	$\Delta pxa1\&2$ -9LHP
Yield (mg/L)	0.57 ± 0.09	0.50 ± 0.07	0.48 ± 0.02	1.21 ± 0.05
Carbon recovery rate (%)	5.80 ± 2.01	5.11 ± 1.63	4.95 ± 0.41	12.40 ± 0.05

4.4 Conclusions

Since the emergence of “metabolic engineering”, its potential in producing unnatural specialty chemicals through genetic and metabolic modification has been extensively explored, especially for the discovery of petroleum-replacing biofuels (82). Among various reported biofuels, hydrocarbons, with high energy density as well as compatibility with current energy storage, transportation and utilization system, stand out as promising petroleum substitutes. While production of short-chained (83-86) and long-chained (19) hydrocarbons have been explored, biofuels and precursors in the medium chain range were seldom reported. In this study, we introduced the hydroperoxide pathway which utilizes linoleic acid as substrate to produce 3(Z)-nonenal, a promising medium-chained hydrocarbon precursor.

Previous studies have discovered that in yeast cells, peroxisomes are the only location where β -oxidation takes place. LCFAs were first activated in cytosol and then transported/translocated into peroxisomes by the ABC transporter, Pxa1/Pxa2 (87). Disruption of either Pxa1 or Pxa2 leads to latency of LCFA β -oxidation while disrupting both genes exhibited a similar phenotype.

In order to study the resulting overall proteins alteration inside the yeast strains, to further characterize the biotransformation potentials, we conducted proteomics by LC-MS/MS techniques. The LC-MS/MS results indicated that 31 proteins displayed significantly different levels in the four functional strains. We categorized the proteins according to their catalyzing functions and the pathways they are involved in: galactose metabolism, glycolysis, TCA cycle, ATP synthesis, amino-acid metabolism, protein biosynthesis, heat shock proteins and others

unknown. After analysis and discussion, the whole-cell catalyst $\Delta pxa1\&2$ -9LHP may display highest biotransformation efficiency.

Subsequently, biotransformation was performed on the functional strains and control strains. While resting cells showed good biotransformation activities, growing cells did not produce detectable amount of 3(Z)-nonenal. Possible explanation is that growing cells were more active in cell divisions rather than performing catalyzing reactions. Furthermore, with the presence of galactose in the culture, which is the preferred carbon source, growing cells would be less likely to take linoleic acid from the medium.

As in the biotransformation cultures, linoleic acid was the sole carbon source. Certain flux ratio would be degraded and generate energy to support the living activities of the cells, including 3(Z)-nonenal biotransformation. Functional strains WT-9LHP, $\Delta pxa1$ -9LHP and $\Delta pxa2$ -9LHP were equivalent in biotransformation efficiency, while the $\Delta pxa1\&2$ -9LHP strain showed a two-fold higher biotransformation, with an efficiency of up to $12.40 \pm 0.05\%$. The biotransformation results were consistent with our expectations from the proteomics analysis results.

In this study, functional strains constructed from single deletion strains displayed similar biotransformation efficiency as that constructed from the wild type strain. The significantly higher biotransformation efficiency of functional strain $\Delta pxa1\&2$ -9LHP indicated that the combination of the two mutations would influence the flux of absorbed linoleic acid and thus result in the retention of more linoleic acid in cytosol to be degraded through the exogenous hydroperoxide

pathway.

In conclusion, we have demonstrated a yeast-based whole-cell biocatalyst capable of transforming PUFAs into medium-chained aldehyde, a medium-chained biofuel precursor. The comparative proteomics analysis offered an approach to study the protein expression levels in the cells and thus potential for further analysis. This study builds the foundations in our future direction to synthesize medium-chained hydrocarbons through metabolic engineering approaches.

4.5 Future directions

This work focused mainly on the production of medium-chained aldehyde production. The raw material for biotransformation used in this research was supplied by the culture medium, which is not economical when functional strains are applied into larger scale of production in the future. Therefore, upstream steps that supply linoleic acid need to be addressed future research.

The source of the raw material is the upstream issue in the use of heterogeneous gene expression. The biosynthesis of FAs is typically initiated by the attachment of acetyl-CoA and *S. cerevisiae* consisted mostly of C₁₆ and C₁₈ FAs, and linoleic acid is one component in the yeast FA pool (88). In this case, accumulation of acetyl-CoA is quite promising to increase the substrate level (89). Genetic modification to yeast β -oxidation has achieved up-regulated MCFA production (90).

Furthermore, work has been reported about the modification of the FA pool in *E.coli* (91). Similar genetic engineering methodology can be applied to *S. cerevisiae*. The development of free FAs as a pool for alkane biosynthesis would

provide a more tractable substrate for biofuel molecules.

The second proposal is to introduce the pathway into oleaginous microorganisms whose higher levels of LCFAs *per se* could support the novel pathway. The potential of oleaginous microorganisms, which accumulate lipid at least 20% of their biomass, including microalgae, bacillus, fungi, for biodiesel production has been studied and discussed (92, 93). Furthermore, the FA biosynthesis pathways as well as the composition of the lipid content in these microorganisms has been analyzed (94, 95).

Among them, research on *Y. lipolytica* showed that this species contains an average level of lipid of 40% of the dry cell weight and the biomass is a rich source of unsaturated FA including linoleic acid, which is the key substrate of our research (96). Moreover, *Y. lipolytica* is the only yeast for which the complete genome is available and where efficient genetic tools have been developed.

In 2004, Santiago-Gomez *et al.* (97, 98) first expressed HPL from green bell pepper in *Y. lipolytica*. Hydroperoxide of linoleic acid was prepared and added into the culture medium. The HPL activity and aldehyde production were assessed and the C6-aldehyde production was as high as 350 mg/L.

In 2013, the LOX and HPL used here were cloned into *Y. lipolytica*. (16). *Y. lipolytica* expressing only one soybean LOX was sufficient to produce pentane. The second step of the reaction was supposed to be spontaneous or catalyzed by LOX. The researchers further improved the yield by altering the C/N supply to shunt carbon flux combined with disrupting one gene involved in β -oxidation and

lipid accumulation. The yield was as much as 4.98 mg/L, which is very promising and impressive.

The LOX and HPL used in this study which mainly focus on the 9-carbon atom molecule and produce 3(Z)-nonenal, in medium-chained range, has not been studied and reported. It is promising to use *Y. lipolytica* as a model organism for lipid accumulation and remobilization. In the future, we will try to further up-regulate the level of aldehyde production.

CHAPTER 5

METABOLICALLY ENGINEERED YEAST CELLS AND MEDIUM-CHAINED HYDROCARBON BIOFUEL SYNTHESIS (PRELIMINARY)

5.1 Introduction

Novel strategies for fossil fuel substitution are being widely explored of late. Researchers have achieved inspiring progress in modifying various microorganisms for the production of biofuel and biofuel intermediates (5, 99).

Alka(e)nes are energy-rich hydrophobic compounds currently being utilized in industry and transportation. The main components of aviation fuels are linear and branched alkanes and cycloalkanes with a typical carbon chain-length distribution of C₆-C₁₆ (4).

The product of interest in last chapter, the fatty aldehyde with a chain length of nine, as the precursor, could be converted into an alka(e)ne, as a potential aviation fuel substitute. The critical enzyme is the non-heme diiron enzyme, cyanobacterial ADC. The alkane biosynthesis pathway was isolated from cyanobacteria. Fatty

acyl-ACPs diverted from fatty aldehydes were diverted into hydrocarbons through decarbonylation.

The catalyzing mechanism of ADC was studied and analyzed, and a model established (100-104). ADC was first revealed to be a non-heme dinuclear iron oxygenase (103). Hydrolysis by ADC is iron-dependent and requires reducing system. The kinetics and necessary cofactors were also discussed. Furthermore, ADC was reported to be oxygen dependent and displayed a faster rate under anaerobic conditions (104).

However, Li *et al.* (105) studied the mechanism of the reaction, proposed that in this reaction the enzyme should be renamed as aldehyde-deformylating oxygenase (ADO). The researchers detected the enzyme activity in the presence and absence of O₂ and ¹⁸O₂ and H₂¹⁸O isotope-tracer and detected the HCO₂⁻ product. The flux of the O-atom was identified by trace isotope analysis. The nature of the reaction was established as oxygenative aldehyde cleavage to formate and alkane.

ADC has been introduced into *E.coli* to produce linear and branched-chain alka(e)nes. After discovery of the alkane biosynthesis pathway in cyanobacteria, Schirmer *et al.* (19) expressed the operons in *E.coli* and a C₁₃ to C₁₇ mixture of alka(e)nes was detected. Howard *et al.*(91) also designed a petroleum-replica hydrocarbon producing pathway in *E.coli* by combining alkane biosynthesis, branched-chain FA production and FA chain length alteration. Furthermore, a pathway producing alkanes with the chain length of C₁₀-C₁₂ was proposed (106). Akhtar *et al.*(107) combined one carboxylic acid reductase (CAR) from *Mycobacterium marinum* and either aldehyde reductase (AHR) or ADC and the catalytic conversion of FAs to fatty alcohols (C₈-C₁₆) or fatty alkanes (C₇-C₁₅) was

reconstituted *in vitro*.

In this chapter, after heterogenous expression of LOX and HPL and successful synthesis of medium-chained aldehyde (as in Chapter 4), one ADC would be expressed in order to convert aldehyde into alka(e)nes. The proposed pathway would be promising to convert renewable raw materials to hydrocarbon biofuels. This would lay the foundation for further development in exploiting bio-techniques toward low-cost renewable transportation fuel production.

The workflow is therefore proposed as shown in Fig. 5.1 below.

Fig. 5.1 Overall approaches to biosynthesize medium-chained biofuel.

5.2 Experiment procedure

In this chapter, we studied hydrocarbon biosynthesis with *S. cerevisiae* as the producing host. The LHPA recombinant plasmid was used to construct and the functional strains as in Chapter 4. The exogenous genes in *S. cerevisiae* were co-expressed and the biotransformation determination was carried out to assess the activities of the whole-cell based catalysis.

The studies described in this chapter are preliminary research. More exploration and optimization is necessary in the future research.

5.3 Results and discussions

5.3.1 Construction of recombinant plasmid

The ADC gene was fused into the 9LHP plasmid and the size of the recombinant plasmid 9LHPA was 13117 bp. Double digestions and electrophoresis were carried out to confirm as shown in Fig. 5.2 (Page 91).

Then recombinant plasmid 9LHPA was then transformed into wild type, single mutant $\Delta pxa1$ and $\Delta pxa2$ and double mutant $\Delta pxa1\&2$ *S. cerevisiae* strains to obtain functional strains WT-9LHPA, $\Delta pxa1$ -9LHPA, $\Delta pxa2$ -9LHPA, $\Delta pxa1\&2$ -9LHPA, respectively. All functional strains were confirmed by colony PCR (data not shown) and grown on YPD and YNB-LEU selective minimal media.

Fig. 5.2 Two different double digestion combination results of recombinant plasmid 9LHPA(13117bp) for confirmation: a) 5k and 8k with *SacI* (4627) and *NheI* (9525); b) 8 k and 5 k with *ClaI* (1604) and *NheI* (9525). (GeneRuler 1 kb Plus DNA Ladder, ready-to-use, Fermentas). See descriptions on Page 90.

5.3.2 Construction of functional strains

The recombinant plasmid 9LHPA was transformed into wild type, single mutant $\Delta pxa1$ and $\Delta pxa2$ and double mutant $\Delta pxa1\&2$ *S. cerevisiae* strains to obtain functional strains WT-9LHPA, $\Delta pxa1$ -9LHPA, $\Delta pxa2$ -9LHPA, $\Delta pxa1\&2$ -9LHPA, respectively.

The colony PCR was carried out to confirm the existence of the recombinant plasmid. Fig. 5.3 (Page 92) shows the colony PCR results against ADC gene. The size of the bands was 460 bp. This proved that the transformation of the recombinant plasmid 9LHPA into the strains was successful.

Fig. 5.3 Colony PCR results of the functional strains against ADC gene, size=460 bp. Lane 4: WT-9LHPA. Lane 5: $\Delta pxa1$ -9LHPA. Lane 6: $\Delta pxa1$ -9LHPA. Lane 7: $\Delta pxa1 \& 2$ -9LHPA. See descriptions on Page 91-92.

5.3.3 Growth curve test

The OD values were measured every time of culture. For the time points 0 h, 4 h and 8 h, the OD values were measured without dilution. For the time points 12 h, 16 h, 20 h, 24 h, 28 h, and 32 h, the OD values were measure after 10 times dilution.

Functional strains WT-9LHPA, $\Delta pxa1$ -9LHPA, $\Delta pxa2$ -9LHPA, $\Delta pxa1 \& 2$ -9LHPA were cultured in YNB-LEU medium for growth curve test. The initial OD value was 0.2. The growth curves in Fig. 5.4 on Page 93 (full data shown in Table S3, Page 126) show that all four functional strains grew well. There was no significant difference among the strains.

Fig. 5.4 Growth curve tested of functional strains WT-9LHPA, $\Delta pxa1$ -9LHPA, $\Delta pxa2$ -9LHPA and $\Delta pxa1 \& 2$ -9LHPA. See descriptions on Page 92-93.

5.3.4 SDS-PAGE and Western blot

In order to study whether the ADC gene was effectively expressed, crude protein samples were extracted from functional strains for Western blot. V5 antibody was used as primary antibody. The existence of the ADC-translated protein proved that ADC can be successfully expressed in functional strains (Shown in Fig. 5.5 below). Strains WT-pESC, $\Delta pxa1$ -pESC, $\Delta pxa2$ -pESC and $\Delta pxa1 \& 2$ -pESC were used as controls.

Fig. 5.5 Western blot of ADC (28 kDa) in *S.cerevisiae* control strains and functional strains; Lane 1: WT-pESC. Lane 2: $\Delta pxa1$ -pESC. Lane 3: $\Delta pxa2$ -pESC. Lane 4: $\Delta pxa1\&2$ -pESC. Lane 5: WT-9LHPA. Lane 6: $\Delta pxa1$ -9LHPA. Lane 7: $\Delta pxa2$ -9LHPA. Lane 8: $\Delta pxa1\&2$ -9LHPA. See descriptions on Page 93-94.

5.3.5 GC-MS results

Functional strains WT-9LHPA, $\Delta pxa1$ -9LHPA, $\Delta pxa2$ -9LHPA and $\Delta pxa1\&2$ -LHPA were cultured, collected and prepared as resting cells for 3 days' biotransformation. Gas samples of the headspaces of the cultures were analyzed by GC-MS. GC spectra of functional strains WT-9LHPA, $\Delta pxa1$ -9LHPA, $\Delta pxa2$ -9LHPA and $\Delta pxa1\&2$ -LHPA are shown in Fig. 5.6 (Page 96). Functional strains WT-9LHP, $\Delta pxa1$ -9LHP, $\Delta pxa2$ -9LHP and $\Delta pxa1\&2$ -LHP were also cultured and detected as a comparison, shown in Fig. 5.7 (Page 97).

However, there is no significant difference between Fig. 5.6 and Fig. 5.7. The peak at 8.10 min was identified as 3(Z)-nonenal by comparing to the retention time of authentic standard.

This peak was further confirmed to be 3(Z)-nonenal by MS spectrum (Fig. 5.8, Page 98).

This proved that the ADC we introduced into the *S. cerevisiae* strains, through translated into enzyme, display non-detectable catalyzing activities.

5.4 Conclusions

The product of chapter 4 is medium-chained aldehyde. While short-chained aldehydes like formaldehyde and acetaldehyde are combustible and regarded as potential fuels replacing fossil fuels, medium-chained aldehydes have not been reported in similar areas.

We introduced the ADC gene into the *S.cerevisiae* strains. The ADC gene was ligated to the 9LHP recombinant plasmid and then was transformed into *S.cerevisiae* wild type and mutants.

Western blot result indicated the existence of the ADC enzyme and proved that the expression of the exogenous gene ADC was successful. However, biotransformation test detected by GC-MS displayed non-detectable catalyzing activity which needs to be addressed in future research.

Fig. 5.6 GC spectra. Peaks at retention time of 8.10 min were identified to be 3(Z)-nonenal. Red, WT-LHPA; Green, YKL-LHPA; Black, YPD-LHPA; Blue:Δ*pxa1* & 2-LHPA. See descriptions on Page 94-95.

Fig. 5.7 GC spectra. Peaks at retention time of 8.10 min were identified to be 3(Z)-nonenal. Red, WT-LHP; Green, YKL-LHP; Black, YPD-LHP; Blue: $\Delta pxal$ & 2-LHP. See descriptions on Page 94-95.

Fig. 5.8 MS spectrum of 3(Z)-nonenal: The peak at retention time of 8.10 min. See descriptions on Page 94-95.

5.5 Future directions

The biotransformation test indicated that the catalyzing activity of ADC was non-detectable.

Recent research proved that the reaction catalyzed by ADC is oxygenative aldehyde cleavage and the enzyme should be renamed as ADO (105). The research on the reaction mechanism proved that the presence of O₂ and NAD(P)H was necessary. Insufficient supplement of NAD(P)H may be the possible reason why no detectable hydrocarbon was produced. In the future we need to provide sufficient cofactors for the reaction. Besides, the kinetic properties of ADO have been studied and calculated (108). The high K_M and low k_{cat} could also be the bottleneck of the biotransformation.

Apart from the oxygen and energy supply, other cofactors were considered. Bernard *et al.* (109) studied alkane biosynthesis in plant: The gene CER1 was reported to be responsible for alkane synthesis and enzymatic complex associated with CER3 was found to be functionally related to ADC. Besides, CYTB5s were likely to be cofactor for CER1 and CER3 and were hypothesized to provide reducing components. The coexpression of these three genes in yeast yielded higher levels of alkane. Similar mechanism maybe exist in *S. cerevisiae*, which is worth exploring.

The production of the medium-chained biofuel precursor, aldehyde, was successful. The future directions should include investigation of both the upstream and the downstream parts of the pathway.

Currently, the microbial platforms for hydrocarbon fuels production do not provide the feasibility of larger scale application. The efficiencies achievable currently were far from those required for commercialization or industrialization.

Factors that may influence the efficiency of alkane synthesis pathways and possible bottlenecks have been discussed (4) which provided us with directions to improve current outcome.

First, the selection of microbial host and optimization of culture conditions may affect the efficiency significantly. In last chapter, we proposed to introduce the aldehyde producing pathway into oleagious yeast which contain much higher levels of FAs *in vivo*, including linoleic acid. In this case, the carbon source in the culture would be glucose instead of linoleic acid which is an economic choice and opens possibilities for larger scale production.

Structure-based modification to improved enzyme properties. In this chapter, though the exogenous genes introduced into the producing host was successful and the translated enzymes displayed catalyzing activities to certain extents, the steric configuration or the kinetics properties of the enzymes have not been explored or discussed. Investigation in this area may provided us information such as substrate preference and binding sites which would be evidence and reference for other modifications and optimizations

Last but not least, is overall pathway optimization. The microbial host itself has a systematic balance while the introduction of exogenous components surely would bring “allergy” to the host. Optimization, including amount of the exogenous

enzymes, cofactors, toxicity and so forth, would enable the host to better cope with exogenous factors.

Biomass is a potential substrate for transportation fuels because of its renewable nature. However, the final goal of biofuel production is to generate directly from sunlight and CO₂ instead of biomass. Although biofuel production still needs further investigation and improvement, the concept has been established. Our discovery provided moderate advancement for further exploration and investigation.

REFERENCE

1. **Keasling JD.** 2010. Manufacturing molecules through metabolic engineering. *Science* **330**:1355-1358.
2. **Regalbuto JR.** 2009. Cellulosic biofuels-got gasoline? *Science* **325**:822-824.
3. **Kung Y, Runguphan W, Keasling JD.** 2012. From fields to fuels: recent advances in the microbial production of biofuels. *Acs Synth Biol* **1**:498-513.
4. **Kallio P, Pásztor A, Akhtar MK, Jones PR.** 2014. Renewable jet fuel. *Curr. Opin. Biotechnol.* **26**:50-55.
5. **McEwen JT, Atsumi S.** 2012. Alternative biofuel production in non-natural hosts. *Curr. Opin. Biotechnol.* **23**:744-750.
6. **Stephanopoulos G.** 2007. Challenges in engineering microbes for biofuels production. *Science* **315**:801-804.
7. **Nasir NF, Daud WRW, Kamarudin SK, Yaakob Z.** 2013. Process system engineering in biodiesel production: A review. *Renew Sust Energ Rev* **22**:631-639.
8. **Sims REH, Mabee W, Saddler JN, Taylor M.** 2010. An overview of second generation biofuel technologies. *Bioresour. Technol.* **101**:1570-1580.
9. **Hansen AC, Zhang Q, Lyne PWL.** 2005. Ethanol-diesel fuel blends-a review. *Bioresour. Technol.* **96**:277-285.
10. **Ladygina N, Dedyukhina EG, Vainshtein MB.** 2006. A review on microbial synthesis of hydrocarbons. *Process Biochem.* **41**:1001-1014.

11. **Shen CR, Liao JC.** 2008. Metabolic engineering of *Escherichia coli* for 1-butanol and 1-propanol production via the keto-acid pathways. *Metab. Eng.* **10**:312-320.
12. **Steen EJ, Yisheng K, Bokinsky G, Zhihao H, Schirmer A, McClure A, del Cardayre SB, Keasling JD.** 2010. Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* **463**:559-562.
13. **Atsumi S, Hanai T, Liao JC.** 2008. Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* **451**:86-89.
14. **Tang XL, Feng HX, Chen WN.** 2013. Metabolic engineering for enhanced fatty acids synthesis in *Saccharomyces cerevisiae*. *Metab. Eng.* **16**:95-102.
15. **Chen WN, Tan KY.** 2013. Malonate uptake and metabolism in *Saccharomyces cerevisiae*. *Appl. Biochem. Biotechnol.* **171**:44-62.
16. **Blazeck J, Liu LQ, Knight R, Alper HS.** 2013. Heterologous production of pentane in the oleaginous yeast *Yarrowia lipolytica*. *J. Biotechnol.* **165**:184-194.
17. **Atsumi S, Higashide W, Liao JC.** 2009. Direct photosynthetic recycling of carbon dioxide to isobutyraldehyde. *Nat. Biotechnol.* **27**:1177-1180.
18. **Kaiser BK, Carleton M, Hickman JW, Miller C, Lawson D, Budde M, Warrenner P, Paredes A, Mullapudi S, Navarro P, Cross F, Roberts JM.** 2013. Fatty aldehydes in cyanobacteria are a metabolically flexible precursor for a diversity of biofuel products. *Plos One* **8**:1-11.
19. **Schirmer A, Rude MA, Li XZ, Popova E, del Cardayre SB.** 2010. Microbial biosynthesis of alkanes. *Science* **329**:559-562.

-
20. **Ducat DC, Way JC, Silver PA.** 2011. Engineering cyanobacteria to generate high-value products. *Trends Biotechnol.* **29**:95-103.
 21. **Hettema EH, Tabak HF.** 2000. Transport of fatty acids and metabolites across the peroxisomal membrane. *Bba-Mol Cell Biol L* **1486**:18-27.
 22. **Abumrad N, Harmon C, Ibrahimi A.** 1998. Membrane transport of long-chain fatty acids: evidence for a facilitated process. *J. Lipid Res.* **39**:2309-2318.
 23. **Hiltunen JK, Mursula AM, Rottensteiner H, Wierenga RK, Kastaniotis AJ, Gurvitz A.** 2003. The biochemistry of peroxisomal beta-oxidation in the yeast *Saccharomyces cerevisiae*. *Fems Microbiol Rev* **27**:35-64.
 24. **van Roermund CWT, Waterham HR, Ijlst L, Wanders RJA.** 2003. Fatty acid metabolism in *Saccharomyces cerevisiae*. *Cell Mol Life Sci* **60**:1838-1851.
 25. **Blackstock WP, Weir MP.** 1999. Proteomics: quantitative and physical mapping of cellular proteins. *Trends Biotechnol.* **17**:121-127.
 26. **Anderson NL, Anderson NG.** 1998. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* **19**:1853-1861.
 27. **Wilkins MR, Pasquali C, Appel RD, Ou K, Golaz O, Sanchez J-C, Yan JX, Gooley AA, Hughes G, Humphery-Smith I, Williams KL, Hochstrasser DF.** 1996. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Nat Biotech* **14**:61-65.
 28. **Pandey A, Mann M.** 2000. Proteomics to study genes and genomes. *Nature* **405**:837-846.

29. **Aebersold R, Mann M.** 2003. Mass spectrometry-based proteomics. *Nature* **422**:198-207.
30. **Gstaiger M, Aebersold R.** 2009. Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat Rev Genet* **10**:617-627.
31. **Doerr A.** 2013. Mass spectrometry-based targeted proteomics. *Nat Methods* **10**:23-23.
32. **Binz P-A, Hochstrasser Denis F, Appel Ron D.** 2003. Mass spectrometry-based proteomics: current status and potential use in clinical chemistry, p. 1540-1551, *Clin Chem Lab Med*, vol. 41.
33. **Ong S-E, Mann M.** 2005. Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* **1**:252-262.
34. **Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R.** 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**:994.
35. **Ross PL, Huang YLN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ.** 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **3**:1154-1169.
36. **Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C.** 2001. Proteolytic ¹⁸O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal Chemi* **73**:2836-2842.
37. **Mirgorodskaya OA, Kozmin YP, Titov MI, Körner R, Sönksen CP, Roepstorff P.** 2000. Quantitation of peptides and proteins by matrix-

- assisted laser desorption/ionization mass spectrometry using ^{18}O -labeled internal standards. *Rapid Commun. Mass Spectrom.* **14**:1226-1232.
38. **Wiese S, Reidegeld KA, Meyer HE, Warscheid B.** 2007. Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics* **7**:340-350.
39. **Griffin TJ, Xie HW, Bandhakavi S, Popko J, Mohan A, Carlis JV, Higgins L.** 2007. iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. *J Proteome Res* **6**:4200-4209.
40. **Zieske LR.** 2006. A perspective on the use of iTRAQTM reagent technology for protein complex and profiling studies. *J. Exp. Bot.* **57**:1501-1508.
41. **Shadforth IP, Dunkley TPJ, Lilley KS, Bessant C.** 2005. i-Tracker: For quantitative proteomics using iTRAQTM. *Bmc Genomics* **6**.
42. **Boundy B, Diegel SW, Wright L, Davis SC** 2011, posting date. Biomass energy data book: edition 4. Oak Ridge National Laboratory, U.S. department of energy. [Online.]
43. **Guldener U, Heck S, Fiedler T, Beinhauer J, Hegemann JH.** 1996. A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res.* **24**:2519-2524.
44. **Guldener U, Heinisch J, Koehler GJ, Voss D, Hegemann JH.** 2002. A second set of *loxP* marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res.* **30**:e23.
45. **Gietz RD, Schiestl RH.** 2007. Frozen competent yeast cells that can be transformed with high efficiency using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* **2**:1-4.

46. **Buchhaupt M, Guder JC, Etschmann MMW, Schrader J.** 2012. Synthesis of green note aroma compounds by biotransformation of fatty acids using yeast cells coexpressing lipoxygenase and hydroperoxide lyase. *Appl. Microbiol. Biotechnol.* **93**:159-168.
47. **Shi JH, Feng HX, Lee J, Chen WN.** 2013. Comparative proteomics profile of lipid-cumulating oleaginous yeast: an iTRAQ-coupled 2-D LC-MS/MS analysis. *Plos One* **8**.
48. **Hughes RK, De Domenico S, Santino A.** 2009. Plant cytochrome CYP74 family: biochemical features, endocellular localisation, activation mechanism in plant defence and improvements for industrial applications. *Chembiochem* **10**:1122-1133.
49. **Feussner I, Kuhn H, Wasternack C.** 2001. Lipoxygenase-dependent degradation of storage lipids. *Trends Plant Sci* **6**:268-273.
50. **Noordermeer MA, Veldink GA, Vliegenthart JFG.** 2001. Fatty acid hydroperoxide lyase: A plant cytochrome P450 enzyme involved in wound healing and pest resistance. *Chembiochem* **2**:494-504.
51. **Matsui K, Fukutomi S, Wilkinson J, Hiatt B, Knauff V, Kajwara T.** 2001. Effect of overexpression of fatty acid 9-hydroperoxide lyase in tomatoes (*Lycopersicon esculentum* Mill.). *J. Agric. Food. Chem.* **49**:5418-5424.
52. **Feussner I, Wasternack C.** 2002. The lipoxygenase pathway. *Annu Rev Plant Biol* **53**:275-297.
53. **Santino A, De Paolis A, Gallo A, Quarta A, Casey R, Mita G.** 2003. Biochemical and molecular characterization of hazelnut (*Corylus avellana*) seed lipoxygenases. *Eur. J. Biochem.* **270**:4365-4375.

-
54. **Chen AO, Whitaker JR.** 1986. Purification and characterization of a lipoxygenase from immature English peas. *J. Agric. Food. Chem.* **34**:203-211.
55. **Gardner HW.** 1989. Soybean lipoxygenase-1 enzymically forms both (9*s*)-hydroperoxides and (13*s*)-hydroperoxides from linoleic-acid by a pH-dependent mechanism. *Biochim. Biophys. Acta* **1001**:274-281.
56. **Georgalaki MD, Bachmann A, Sotiroudis TG, Xenakis A, Porzel A, Feussner I.** 1998. Characterization of a 13-lipoxygenase from virgin olive oil and oil bodies of olive endosperms. *Fett-Lipid* **100**:554-560.
57. **Mita G, Gallo A, Greco V, Zasiura C, Casey R, Zacheo G, Santino A.** 2001. Molecular cloning and biochemical characterization of a lipoxygenase in almond (*Prunus dulcis*) seed. *Eur. J. Biochem.* **268**:1500-1507.
58. **Santino A, Iannaccone R, Hughes R, Casey R, Mita G.** 2005. Cloning and characterisation of an almond 9-lipoxygenase expressed early during seed development. *Plant Sci* **168**:699-706.
59. **Matsui K, Toyota H, Kajiwara T, Kakuno T, Hatanaka A.** 1991. Fatty-acid hydroperoxide cleaving enzyme, hydroperoxide lyase, from tea leaves. *Phytochemistry* **30**:2109-2113.
60. **Shibata Y, Matsui K, Kajiwara T, Hatanaka A.** 1995. Purification and properties of fatty-acid hydroperoxide lyase from green bell pepper fruits. *Plant Cell Physiol* **36**:147-156.
61. **Matoba T, Hidaka H, Kitamura K, Kaizuma N, Kito M.** 1985. Contribution of hydroperoxide lyase activity to normal-hexanal formation in soybean. *J. Agric. Food. Chem.* **33**:856-858.
-

62. **Noordermeer MA, Veldink GA, Vliegthart JFG.** 1999. Alfalfa contains substantial 9-hydroperoxide lyase activity and a 3Z : 2E-enal isomerase. *FEBS Lett.* **443**:201-204.
63. **Tijet N, Schneider C, Muller BL, Brash AR.** 2001. Biogenesis of volatile aldehydes from fatty acid hydroperoxides: molecular cloning of a hydroperoxide lyase (CYP74C) with specificity for both the 9-and 13-hydroperoxides of linoleic and linolenic acids. *Arch. Biochem. Biophys.* **386**:281-289.
64. **Matsui K, Ujita C, Fujimoto S, Wilkinson J, Hiatt B, Knauf V, Kajiwara T, Feussner I.** 2000. Fatty acid 9-and 13-hydroperoxide lyases from cucumber. *FEBS Lett.* **481**:183-188.
65. **Mita G, Quarta A, Fasano P, De Paolis A, Di Sansebastiano GP, Perrotta C, Iannacone R, Belfield E, Hughes R, Tsesmetzis N, Casey R, Santino A.** 2005. Molecular cloning and characterization of an almond 9-hydroperoxide lyase, a new CYP74 targeted to lipid bodies. *J. Exp. Bot.* **56**:2321-2333.
66. **Holden HM, Rayment I, Thoden JB.** 2003. Structure and function of enzymes of the Leloir pathway for galactose metabolism. *J. Biol. Chem.* **278**:43885-43888.
67. **Sellick CA, Campbell RN, Reece RJ.** 2008. Chapter 3 Galactose metabolism in yeast-structure and regulation of the Leloir Pathway enzymes and the genes encoding them, p. 111-150. *In* Kwang WJ (ed.), *International Review of Cell and Molecular Biology*, vol. Volume 269. Academic Press.

-
68. **Johnston M, Flick JS, Pexton T.** 1994. Multiple mechanisms provide rapid and stringent glucose repression of GAL gene-expression in *Saccharomyces cerevisiae*. *Mol Cell Biol* **14**:3834-3841.
69. **Voet D, Voet JG, Pratt CW.** 2002. Fundamentals of biochemistry. New York : Wiley, c2002.
70. **Velours J, Paumard P, Soubannier V, Spannagel C, Vaillier J, Arselin G, Graves PV.** 2000. Organisation of the yeast ATP synthase F₀: a study based on cysteine mutants, thiol modification and cross-linking reagents. *Bba-Bioenergetics* **1458**:443-456.
71. **Devenish RJ, Prescott M, Roucou X, Nagley P.** 2000. Insights into ATP synthase assembly and function through the molecular genetic manipulation of subunits of the yeast mitochondrial enzyme complex. *Bba-Bioenergetics* **1458**:428-442.
72. **Bender DA.** 2012. The branched-chain amino acids: leucine, isoleucine and valine, p. 279-303, *Amino Acid Metabolism*. John Wiley & Sons, Ltd.
73. **Whitfield CD, Steers EJ, Weissbach H.** 1970. Purification and properties of 5-methyltetrahydropteroyltriglutamate-homocysteine transmethylase. *J. Biol. Chem.* **245**:390-401.
74. **Pejchal R, Ludwig ML.** 2005. Cobalamin-independent methionine synthase (MetE): A face-to-face double barrel that evolved by gene duplication. *Plos Biol* **3**:254-265.
75. **Dickinson JR, Salgado LEJ, Hewlins MJE.** 2003. The catabolism of amino acids to long chain and complex alcohols in *Saccharomyces cerevisiae*. *J. Biol. Chem.* **278**:8028-8034.
-

76. **Dominguez D, Altmann M, Benz J, Baumann U, Trachsel H.** 1999. Interaction of translation initiation factor eIF4G with eIF4A in the yeast *Saccharomyces cerevisiae*. J. Biol. Chem. **274**:26720-26726.
77. **Neff CL, Sachs AB.** 1999. Eukaryotic translation initiation factors 4G and 4A from *Saccharomyces cerevisiae* interact physically and functionally. Mol Cell Biol **19**:5557-5564.
78. **Chuang SM, Chen L, Lambertson D, Anand M, Kinzy TG, Madura M.** 2005. Proteasome-mediated degradation of cotranslationally damaged proteins involves translation elongation factor 1A. Mol Cell Biol **25**:403-413.
79. **Stelzl U, Connell S, Nierhaus KH, Wittmann-Liebold B.** 2001. Ribosomal proteins: role in ribosomal functions, eLS. John Wiley & Sons, Ltd.
80. **Sturtz LA, Diekert K, Jensen LT, Lill R, Culotta VC.** 2001. A fraction of yeast Cu,Zn-superoxide dismutase and its metallochaperone, CCS, localize to the intermembrane space of mitochondria: a physiological role for SOD1 in guarding against mitochondrial oxidative damage J. Biol. Chem. **276**:38084-38089.
81. **Julsing MK, Kuhn D, Schmid A, Buhler B.** 2012. Resting cells of recombinant *E. coli* show high epoxidation yields on energy source and high sensitivity to product inhibition. Biotechnol. Bioeng. **109**:1109-1119.
82. **Keasling JD.** 2012. Synthetic biology and the development of tools for metabolic engineering. Metab. Eng. **14**:189-195.

-
83. **Smith KM, Cho KM, Liao JC.** 2010. Engineering *Corynebacterium glutamicum* for isobutanol production. Appl. Microbiol. Biotechnol. **87**:1045-1055.
84. **Lan EI, Liao JC.** 2011. Metabolic engineering of cyanobacteria for 1-butanol production from carbon dioxide. Metab. Eng. **13**:353-363.
85. **Baez A, Cho KM, Liao JC.** 2011. High-flux isobutanol production using engineered *Escherichia coli*: a bioreactor study with in situ product removal. Appl. Microbiol. Biotechnol. **90**:1681-1690.
86. **Steen E, Chan R, Prasad N, Myers S, Petzold C, Redding A, Ouellet M, Keasling J.** 2008. Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. Microb Cell Fact **7**:36.
87. **Hettema EH, vanRoermund CWT, Distel B, vandenBerg M, Vilela C, RodriguesPousada C, Wanders RJA, Tabak HF.** 1996. The ABC transporter proteins Pat1 and Pat2 are required for import of long-chain fatty acids into peroxisomes of *Saccharomyces cerevisiae*. EMBO J. **15**:3813-3822.
88. **Tehlivets O, Scheuringer K, Kohlwein SD.** 2007. Fatty acid synthesis and elongation in yeast. Bba-Mol Cell Biol L **1771**:255-270.
89. **Chen.** 2014. Enhancement of free fatty acid production in *Saccharomyces cerevisiae* by control of fatty acyl-CoA metabolism.
90. **Chen LW, Zhang JH, Chen WN.** 2014. Engineering the *Saccharomyces cerevisiae* beta-oxidation pathway to increase medium chain fatty acid production as potential biofuel. Plos One **9**.
91. **Howard TP, Middelhaufe S, Moore K, Edner C, Kolak DM, Taylor GN, Parker DA, Lee R, Smirnoff N, Aves SJ, Love J.** 2013. Synthesis of

- customized petroleum-replica fuel molecules by targeted modification of free fatty acid pools in *Escherichia coli*. P Natl Acad Sci USA **110**:7636-7641.
92. **Meng X, Yang J, Xu X, Zhang L, Nie Q, Xian M.** 2009. Biodiesel production from oleaginous microorganisms. Renew Energ **34**:1-5.
93. **Papanikolaou S, Chevalot I, Komaitis M, Marc I, Aggelis G.** 2002. Single cell oil production by *Yarrowia lipolytica* growing on an industrial derivative of animal fat in batch cultures. Appl. Microbiol. Biotechnol. **58**:308-312.
94. **Ratledge C.** 2004. Fatty acid biosynthesis in microorganisms being used for single cell oil production. Biochimie **86**:807-815.
95. **Guil-Guerrero J, Navarro-Juárez R, López-Martinez J, Campra-Madrid P, Reboloso-Fuentes M.** 2004. Functional properties of the biomass of three microalgal species. J. Food Eng. **65**:511-517.
96. **Nicaud JM.** 2012. *Yarrowia lipolytica*. Yeast **29**:409-418.
97. **Bourel G, Nicaud JM, Nthangeni B, Santiago-Gomez P, Belin JM, Husson F.** 2004. Fatty acid hydroperoxide lyase of green bell pepper: cloning in *Yarrowia lipolytica* and biogenesis of volatile aldehydes. Enzyme Microb. Technol. **35**:293-299.
98. **Santiago-Gomez MP, Thanh HT, De Coninck J, Cachon R, Kermasha S, Belin JM, Gervais P, Husson F.** 2009. Modeling hexanal production in oxido-reducing conditions by the yeast *Yarrowia lipolytica*. Process Biochem. **44**:1013-1018.
99. **Gronenberg LS, Marcheschi RJ, Liao JC.** 2013. Next generation biofuel engineering in prokaryotes. Curr. Opin. Chem. Biol. **17**:462-471.

-
100. **Warui DM, Li N, Norgaard H, Krebs C, Bollinger JM, Booker SJ.** 2011. Detection of formate, rather than carbon monoxide, as the stoichiometric coproduct in conversion of fatty aldehydes to alkanes by a cyanobacterial aldehyde decarbonylase. *J. Am. Chem. Soc.* **133**:3316-3319.
101. **Li N, Norgaard H, Warui DM, Booker SJ, Krebs C, Bollinger JM.** 2011. Conversion of fatty aldehydes to alka(e)nes and formate by a cyanobacterial aldehyde decarbonylase: cryptic redox by an unusual dimetal oxygenase. *J. Am. Chem. Soc.* **133**:6158-6161.
102. **Paul B, Das D, Ellington B, Marsh ENG.** 2013. Probing the mechanism of cyanobacterial aldehyde decarbonylase using a cyclopropyl aldehyde. *J. Am. Chem. Soc.* **135**:5234-5237.
103. **Das D, Eser BE, Han J, Sciore A, Marsh ENG.** 2011. Oxygen-independent decarbonylation of aldehydes by cyanobacterial aldehyde decarbonylase: a new reaction of diiron enzymes. *Angew. Chem.-Int. Edit.* **50**:7148-7152.
104. **Eser BE, Das D, Han J, Jones PR, Marsh ENG.** 2011. Oxygen-independent alkane formation by non-heme iron-dependent cyanobacterial aldehyde decarbonylase: investigation of kinetics and requirement for an external electron donor. *Biochemistry-Us* **50**:10743-10750.
105. **Li N, Chang W-c, Warui DM, Booker SJ, Krebs C, Bollinger JM.** 2012. Evidence for only oxygenative cleavage of aldehydes to alk(a/e)nes and formate by cyanobacterial aldehyde decarbonylases. *Biochemistry-Us* **51**:7908-7916.
106. **Choi YJ, Lee SY.** 2013. Microbial production of short-chain alkanes. *Nature* **502**:571-574.
-

107. **Akhtar MK, Turner NJ, Jones PR.** 2013. Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities. *P Natl Acad Sci USA* **110**:87-92.
108. **Khara B, Menon N, Levy C, Mansell D, Das D, Marsh ENG, Leys D, Scrutton NS.** 2013. Production of propane and other short-chain alkanes by structure-based engineering of ligand specificity in aldehyde-deformylating oxygenase. *ChemBioChem* **14**:1204-1208.
109. **Bernard A, Domergue F, Pascal S, Jetter R, Renne C, Faure J-D, Haslam RP, Napier JA, Lessire R, Joubès J.** 2012. Reconstitution of plant alkane biosynthesis in yeast demonstrates that *Arabidopsis* ECERIFERUM1 and ECERIFERUM3 are core components of a very-long-chain alkane synthesis complex. *The Plant Cell Online* **24**:3106-3118.

SUPPLEMENT

Table S1 Data of growth curve Fig. 4.8 (Page 62)

WT-pESC					
t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.512	0.622	0.609	0.581	0.060
8	1.997	2.089	2.101	2.062	0.057
12	5.660	5.910	5.830	5.800	0.128
16	7.530	7.800	7.880	7.737	0.183
20	9.670	10.200	10.240	10.037	0.318
24	9.480	10.110	8.920	9.503	0.595
28	9.870	10.230	10.250	10.117	0.214
32	10.840	11.560	10.720	11.040	0.454

<i>Δpxa1</i>- pESC					
t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.488	0.503	0.538	0.510	0.026
8	1.802	1.840	1.952	1.865	0.078
12	5.110	5.240	5.260	5.203	0.081
16	7.240	7.720	8.070	7.677	0.417
20	9.780	10.040	10.100	9.930	0.170
24	9.680	10.140	10.490	10.103	0.406
28	10.010	10.630	10.850	10.497	0.436
32	10.120	10.420	10.350	10.297	0.157

<i>Δpxa2</i>- pESC					
t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.544	0.530	0.517	0.530	0.014
8	1.975	1.949	1.937	1.954	0.019
12	5.540	5.390	5.280	5.403	0.131
16	8.300	8.210	8.260	8.257	0.045

20	9.860	9.650	9.580	9.697	0.146
24	9.880	9.620	9.320	9.607	0.280
28	10.220	10.250	10.160	10.210	0.046
32	10.110	10.130	10.080	10.107	0.025

Δpxa1&2- pESC

t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.634	0.623	0.635	0.631	0.007
8	2.113	2.079	2.130	2.107	0.026
12	6.020	5.890	6.130	6.013	0.120
16	8.450	8.490	8.780	8.573	0.180
20	9.980	10.100	9.930	10.003	0.087
24	10.230	10.340	10.620	10.397	0.201
28	10.880	10.850	10.960	10.897	0.057
32	11.070	10.130	11.800	11.000	0.837

WT-9LHP

t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.428	0.460	0.457	0.448	0.018
8	1.612	1.592	1.620	1.608	0.014
12	3.990	4.180	4.030	4.067	0.100
16	6.480	6.530	6.680	6.563	0.104
20	8.020	8.100	8.300	8.140	0.144
24	7.670	7.710	8.000	7.793	0.180
28	8.050	8.330	8.470	8.283	0.214
32	8.020	8.290	8.420	8.243	0.204

Δpxa1-9LHP

t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.463	0.429	0.430	0.441	0.019
8	1.411	1.437	1.444	1.431	0.017
12	4.230	4.310	4.270	4.270	0.040
16	6.670	6.720	6.820	6.737	0.076

20	7.970	8.130	8.060	8.053	0.080
24	8.310	8.120	8.040	8.157	0.139
28	8.350	8.480	8.410	8.413	0.065
32	8.830	9.230	8.960	9.007	0.204

Δpxa2-9LHP

t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.483	0.491	0.498	0.491	0.008
8	1.498	1.515	1.493	1.502	0.012
12	4.210	4.260	4.270	4.247	0.032
16	6.080	6.180	6.240	6.167	0.081
20	7.830	7.970	7.920	7.907	0.071
24	8.040	7.980	8.140	8.053	0.081
28	7.910	8.120	7.690	7.907	0.215
32	8.760	8.640	9.280	8.893	0.340

Δpxa1&2-9LHP

t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.482	0.513	0.520	0.505	0.020
8	2.033	2.247	1.997	2.092	0.135
12	4.730	5.030	4.883	4.881	0.150
16	6.870	7.110	6.820	6.933	0.155
20	8.130	8.340	8.170	8.213	0.112
24	8.340	8.670	8.510	8.507	0.165
28	8.930	9.470	8.870	9.090	0.330
32	9.720	10.020	9.940	9.893	0.155

Fig. S1 Total intensity chromatogram results of peptides eluted by gradient concentrations of APS. See descriptions on Page 64.

Fig. S2 LC-MS qualification results of representative peptide fragmentation spectrum of glucose-6-phosphate isomerase. See descriptions on Page 64.

Agilent Spectrum Mill - Protein/Peptide Summary									
Spectrum Mill	Summary Settings	Autovalidation	Easy MS/MS	MS/MS Search	Spectrum Summary	Build TIC	Tool Belt	Help	
1	MSNNFTNFK	LATELPAWSK	LQKIYESQ GK	TL SVKQEFQK	DAKRFEKLNK	TFTNYDGSKI	LFDYSKNLVN	DEIIAALIEL	80
81	AK EANVTGLR	DAMFKGEHIN	STEDRAVYHV	ALRNRANKPM	YVDGVNVAPE	VDSVLKHMKE	FSEQVRSGEW	KGYTGKKITD	160
161	VVNIGIGGSD	LGPVMVTEAL	KHYAGVLDVH	FVSNIDGTHI	AETLKVVDP E	TTLFLIASKT	FTTAETITNA	NTAKNWFLSK	240
241	TGNDP SHIAK	HFAALSTNET	EVAKFGIDTK	NMFGFESWVG	GRYSVWSAIG	LSVALYIGYD	NFEAFLKGAE	AVDNHFTQTP	320
321	LEDNIPLLGG	LLSVWYNNFF	GAQTHLVAPF	DQYLHRFPAY	LQQLSMESNG	KSVTRGNVFT	DYSTGSILFG	EPATNAQHSF	400
401	FQLVHQGTKL	IPSDFILAAQ	SHNPIENKLH	QKMLASNFFA	QAEALMVGKD	EEQVKAEGAT	GGLVPHKVFS	GNRPPTSILA	480
481	QKITPATLGA	LIAYYEHVTF	TEGAIWNINS	FDQWGV ELGK	VLAKVIGKEL	DNSSTISTHD	ASTNGLINQF	KEWM	554

The matched peptides cover 8% (45/554 AA's) of the protein.

Protein Name: **Glucose-6-phosphate isomerase OS=Saccharomyces cerevisiae GN=PGI1 PE=1 SV=3**
 Species: **YEAST**
 SwissProt Accession # **P12709**
 MS Digest Index # **128672**
 Masses are:
 pl of Protein: **6.00**
 Protein MW: **61299.5 Da**
 Amino Acid Composition: **A48 D25 E35 F34 G41 H18 I30 K40 L47 M10 N37 P16 Q20 R10 S38 T42 V38 W9 Y16**

MS Digest - Agilent Spectrum Mill Rev. 3.3.084

Fig. S3 Peptide summary of glucose-6-phosphate isomerase. See descriptions on Page 64.

Agilent Spectrum Mill - Protein/Peptide Summary							
Spectrum Mill	Summary Settings	Autovalidation	Easy MS/MS	MS/MS Search	Spectrum Summary	Build TIC	Tool Belt Help
MS Digest Results							
Protein Name: Glucose-6-phosphate isomerase OS=Saccharomyces cerevisiae GN=PGI1 PE=1 SV=3 Species: YEAST SwissProt Accession # P12709 MS Digest Index # 128672 Fixed Modifications: MMTS iTRAQ8plex Variable Modifications: (None) Masses are: Monoisotopic pI of Protein: 6.00 Protein MW: 61299.5 Da Amino Acid Composition: A48 D25 E35 F34 G41 H18 I30 K40 L47 M10 N37 P16 Q20 R10 S38 T42 V38 W9 Y16							
Digest Used	Max. # Missed Cleavages	Cysteines Modified by	Peptide N terminus		Peptide C terminus		
Trypsin	2	MMTS	iTRAQ_8plex (H25 C7 N3 O3 c7 n1)		Free Acid (O H)		
1	11	21	31	41	51	61	71
MSNNSFTNFK	LATELPAWSK	LQKIYESQGR	TLSVKQEFQK	DAKRFEKLNK	TFTNYDGSKI	LPDYSKNLVN	DEIIAALIEL
81	91	101	111	121	131	141	151
AKEANVTGLR	DAMFKGEHIN	STEDRAVYHV	ALRNANKPM	YVDGVNVAPE	VDSVLKHMKE	FSEQVRSGEW	KGYTGKKITD
161	171	181	191	201	211	221	231
VVNIGIGSD	LGPVMVTEAL	KHYAGVLDVH	FVSNIDGTHI	AETLKVVDP	TTLFLIASKT	FTTAETITNA	NTAKNWFLSK
241	251	261	271	281	291	301	311
TGNDFSHIAK	HFAALSTNET	EVAKFGIDTK	NMFGFESWVG	GRYSVWSAIG	LSVALYIGYD	NFEAPLKGAE	AVDNHFTQTP
321	331	341	351	361	371	381	391
LEDNIPLGG	LLSVWYNNFF	GAQTHLVAPF	DQYLHFFPAY	LQQLSMESNG	KSVTRGNVFT	DYSTGSILFG	EPATNAQHSF
401	411	421	431	441	451	461	471
FQLVHQGTKL	IPSEFILAAQ	SHNPIENKLH	QKMLASNFFA	QAEALMVGKD	EEQVKAEGAT	GGLVPHKVFS	GNRPTTSILA
481	491	501	511	521	531	541	551
QKITPATLGA	LIAYYEHVTF	TEGAIWNINS	FDQMGVELGK	VLAQVIGKEL	DNSSTISTHD	ASTNGLINQF	KEWM
The matched peptides cover 0% (0/554 AA's) of the protein.							
Start	End	MH ⁺	Sequence				
152	156	1133.6775	(K)GYTGK(K)				
31	35	1155.7557	(K)TLSVK(Q)				
83	90	1163.6686	(K)EANVTGLR(D)				

Fig. S4 MS digest results of glucose-6-phosphate isomerase (part 1). See descriptions on Page 64.

Start	End	MH ⁺	Sequence
152	156	1133.6775	(K)GYTGK(K)
31	35	1155.7557	(K)TLVK(Q)
83	90	1163.6686	(K)EANTGLR(D)
140	146	1198.6369	(K)EFSEQVR(S)
147	151	1214.6989	(R)SGEWK(G)
91	95	1219.6965	(R)DAMFK(G)
106	113	1232.7417	(R)AVYHVALR(N)
36	40	1287.7517	(K)QEFQK(D)
265	270	1288.7721	(K)FGDTK(N)
450	455	1355.7626	(K)DEEQVK(A)
235	240	1402.8303	(K)NWFLSK(T)
24	30	1432.8256	(K)YESQK(T)
96	105	1461.7235	(K)GEHINSTEDR(A)
60	66	1493.8824	(K)ILFDYSK(N)
106	115	1502.8857	(R)AVYHVALRNR(A)
152	157	1565.9778	(K)GYTGKK(I)
51	59	1640.8740	(K)TFTNYDGSK(I)
241	250	1647.9274	(K)TGNDPSHIK(H)
271	282	1690.8313	(K)NMFGEFESWVGGR(Y)
45	50	1691.0618	(R)FEKLNK(T)
11	20	1724.0203	(K)LATELPAWSK(L)
521	528	1740.1874	(K)VLAKVIGK(E)
456	467	1745.0166	(K)AEGATGGLVPHK(V)
1	10	1797.9414	(-)MSNNSFTNFK(L)
41	47	1806.1000	(K)DAKRFEK(L)
44	50	1847.1630	(K)RFEKLNK(T)
137	146	1899.0367	(K)HMKFESEQVR(S)
36	43	1906.1161	(K)QEFQKDAK(R)
147	156	2025.1532	(R)SGEWKGYTGK(K)
83	95	2060.1418	(K)EANTGLRDAMFK(G)
36	44	2062.2172	(K)QEFQKDAKR(F)
140	151	2090.1127	(K)EFSEQVRSGEWK(G)
21	30	2106.2685	(K)LQKYESQK(T)

Fig. S5 MS digest results of glucose-6-phosphate isomerase (part 2). See descriptions on Page 64.

31	40	2120.2842	(K)TLVKQFFQK(D)
251	264	2126.1702	(K)HFAALSTNETEVAK(F)
206	219	2141.2678	(K)VVDPETTLFLASK(T)
220	234	2192.2019	(K)TFTTAETITNANTAK(N)
468	482	2227.3019	(K)VFSGNRPTTSILAQK(I)
24	35	2265.3581	(K)YESQGKTLVK(Q)
48	59	2300.3013	(K)LNKFTFTNYDGSK(I)
357	371	2321.2420	(R)FPAYLQQLSMESNGK(S)
67	82	2347.4056	(K)NLVNDEIIAALIELAK(E)
91	105	2358.1968	(R)DAMFKGEHINSTEDR(A)
96	113	2371.2420	(K)GEHINSTEDRAVYHVALR(N)
11	23	2397.4632	(K)LATELPAWSKLQK(I)
433	449	2436.3239	(K)MLASNFFAQAEALMVGK(D)
147	157	2457.4535	(R)SGEWKGYTGKK(I)
96	115	2641.3860	(K)GEHINSTEDRAVYHVALRNR(A)
265	282	2656.3802	(K)FGIDTKNMFGEFESWVGGR(Y)
410	428	2715.5289	(K)LIPSFILAAQSHNPIENK(L)
235	250	2727.5345	(K)NWFLSKTGNDPSHIK(H)
31	43	2738.6486	(K)TLVKQEFQKDAK(R)
357	375	2764.4912	(R)FPAYLQQLSMESNGKSVTR(G)
450	467	2777.5560	(K)DEEQVKAEGATGGLVPHK(V)
137	151	2790.5124	(K)HMKFEQVRSGEWK(G)
51	66	2811.5331	(K)TFTNYDGSKILFDYSK(N)
140	156	2900.5669	(K)EFSEQVRSGEWKGYTGK(K)
21	35	2938.8011	(K)LQKYESQGKTLVK(Q)
158	181	3006.7005	(K)TDVNIIGIGGSDLGPVMVTEALK(H)
45	59	3008.7126	(R)FEKLNKFTFTNYDGSK(I)
251	270	3091.7191	(K)HFAALSTNETEVAKFGIDTK(N)
529	551	3100.6007	(K)ELDNSSTISTHDASTNGLINQFK(E)
116	136	3157.7694	(R)ANKPMYVDGVNVAPEVDSVLK(H)
67	90	3187.8510	(K)NLVNDEIIAALIELAKEANVTGLR(D)
83	105	3198.6422	(K)EANVTGLRDAMFKGEHINSTEDR(A)
1	20	3198.7384	(-)MSNNSFTNFKLATELPAWSK(L)
24	40	3229.8866	(K)YESQGKTLVKQFFQK(D)

Fig. S6 MS digest results of glucose-6-phosphate isomerase (part 3). See descriptions on Page 64.

182 205 3244.7574 (K)HYAGVLDVHFVSNIDGTHIAETLK(V)
 429 449 3246.8258 (K)LHQKMLASNFFAQAEALMVGK(D)
 91 113 3267.7153 (R)DAMFKGEHINSTEDRAVYHVALR(N)
 220 240 3271.8089 (K)TFTTAETITNANTAKNWFLSK(T)
 114 136 3427.9134 (R)NRANKPMYVDGVNVAPEVDSVLK(H)
 283 307 3434.8496 (R)YSVWSAIGLSVALYIGYDNFEAFK(G)
 157 181 3439.0008 (K)KITDVVNIGIGGSDLGPVMVTEALK(H)
 241 264 3450.8744 (K)TGNDPSHIAKHFAALSTNETEVAK(F)
 433 455 3468.8634 (K)MLASNFFAQAEALMVGKDEEQVK(A)
 48 66 3470.9605 (K)LNKTFTNYDGSKILFDYSK(N)
 11 30 3507.0656 (K)LATELPAWSKLQKIYESQGK(T)
 60 82 3518.0648 (K)ILFDYSKNLVNDEIIAALIELAK(E)
 410 432 3526.0308 (K)IPSDFILAAQSHNPIENKLHQK(M)
 529 554 3546.7631 (K)ELDNSSTISTHDASTNGLNQFKEWM(-)
 456 482 3649.0952 (K)AEGATGGLVPHKVFSGNRPTTSILAQK(I)
 525 551 3802.0749 (K)VGKELDNSSTISTHDASTNGLNQFK(E)
 116 139 3858.1691 (R)ANKPMYVDGVNVAPEVDSVLKHKM(E)
 1 23 3872.1814 (-)MSNNSFTNFKLATELPAWSKLQK(I)
 206 234 4010.2464 (K)VDPETTLFLIASKTFTTAETITNANTAK(N)
 67 95 4084.3243 (K)NLVNDEIIAALIELAKEANYTGLRDAMFK(G)
 114 139 4128.3131 (R)NRANKPMYVDGVNVAPEVDSVLKHKM(E)
 525 554 4248.2373 (K)VGKELDNSSTISTHDASTNGLNQFKEWM(-)
 152 181 4249.4551 (K)GYTGKKITDVVNIGIGGSDLGPVMVTEALK(H)
 429 455 4279.3652 (K)LHQKMLASNFFAQAEALMVGKDEEQVK(A)
 376 409 4307.1981 (R)GNVFTDYSTGSILFGEPATNAQHSFFQLVHQGTK(L)
 106 136 4337.4319 (R)AVYHVALNRANKPMYVDGVNVAPEVDSVLK(H)
 60 90 4358.5102 (K)ILFDYSKNLVNDEIIAALIELAKEANYTGLR(D)
 241 270 4416.4233 (K)TGNDPSHIAKHFAALSTNETEVAKFGIDTK(N)
 251 282 4459.3271 (K)HFAALSTNETEVAKFGIDTKNMFGEFESWVGGR(Y)
 521 551 4517.5648 (K)VLAKVGKELDNSSTISTHDASTNGLNQFK(E)
 235 264 4530.4815 (K)NWFLSKTGNDPSHIAKHFAALSTNETEVAK(F)
 220 250 4596.5131 (K)TFTTAETITNANTAKNWFLSKTGNDPSHIAK(H)
 450 482 4681.6347 (K)DEEQVKAEGATGGLVPHKVFSGNRPTTSILAQK(I)
 116 146 4733.5828 (R)ANKPMYVDGVNVAPEVDSVLKHKMKFESEQVR(S)

Fig. S7 MS digest results of glucose-6-phosphate isomerase (part 4). See descriptions on Page 64.

235 264 4530.4615 (K)NWFLSKTGNDPSHIAK(H)
 220 250 4596.5131 (K)TFTTAETITNANTAKNWFLSKTGNDPSHIAK(H)
 450 482 4681.6347 (K)DEEQVKAEGATGGLVPHKVFSGNRPTTSILAQK(I)
 116 146 4733.5825 (R)ANKPMYVDGVNVAPEVDSVLKHKMEFSEQVR(S)
 372 409 4750.4474 (K)SVTRGNVFTDYSTGSILFGEPATNAQHSFFQLVHQGT(L)
 271 307 4802.4577 (K)NMFGEFESWVGGRYSVWSAIGLSVALYIGYDNFEAFLK(G)
 483 520 4833.5388 (K)TPATLGALIAYYEHVTFTEGAIWNINSFDQWGVLELKG(V)
 51 82 4835.7156 (K)TFTNYDGSILFDYSKNLVNDEIAALIELAK(E)
 433 467 4890.6567 (K)MLASNFFAQAEALMVGKDEEQVKAEGATGGLVPHK(V)
 182 219 5062.8020 (K)HYAGVLDVHFVSNIDGTHIAETLKVDPETTLFLIASK(T)
 206 240 5089.8535 (K)VDPETTLFLIASKTFTTAETITNANTAKNWFLSK(T)
 483 524 5549.0287 (K)TPATLGALIAYYEHVTFTEGAIWNINSFDQWGVLELKGVLAK(V)
 410 449 5639.1315 (K)LIPSDFILAAQSHNPIENKLHQMLASNFFAQAEALMVGK(D)
 265 307 5768.0066 (K)FGIDTKNMFGEFESWVGGRYSVWSAIGLSVALYIGYDNFEAFLK(G)
 308 356 5829.9396 (K)GAEAVDNHFTQTPLEDNIPLLGGLLSVWYNNFFGAQTHLVAPFDQYLHR(F)
 158 205 5928.2348 (K)TDVWNIGGSDLGPMVTEALKHYAGVLDVHFVSNIDGTHIAETLK(V)
 483 528 6250.5030 (K)TPATLGALIAYYEHVTFTEGAIWNINSFDQWGVLELKGVLAKVIGK(E)
 157 205 6360.5351 (K)KTDVWNIGGSDLGPMVTEALKHYAGVLDVHFVSNIDGTHIAETLK(V)
 376 428 6699.5039 (R)GNVFTDYSTGSILFGEPATNAQHSFFQLVHQGT(L)
 468 520 6737.6174 (K)VFSGNRPTTSILAQKITPATLGALIAYYEHVTFTEGAIWNINSFDQWGVLELKG(V)
 357 409 6748.4661 (R)FPAYLQQLSMESNGKSVTRGNVFTDYSTGSILFGEPATNAQHSFFQLVHQGT(L)
 182 234 6931.7807 (K)HYAGVLDVHFVSNIDGTHIAETLKVDPETTLFLIASKTFTTAETITNANTAK(N)
 372 428 7142.7531 (K)SVTRGNVFTDYSTGSILFGEPATNAQHSFFQLVHQGT(L)
 468 524 7453.1074 (K)VFSGNRPTTSILAQKITPATLGALIAYYEHVTFTEGAIWNINSFDQWGVLELKGVLAK(V)
 376 432 7510.0057 (R)GNVFTDYSTGSILFGEPATNAQHSFFQLVHQGT(L)
 158 219 7746.2793 (K)TDVWNIGGSDLGPMVTEALKHYAGVLDVHFVSNIDGTHIAETLKVDPETTLFLIASK(T)
 308 371 7827.9583 (K)GAEAVDNHFTQTPLEDNIPLLGGLLSVWYNNFFGAQTHLVAPFDQYLHRFPAYLQQLSMESNGK(S)
 456 520 8159.4108 (K)AEGATGGLVPHKVFSGNRPTTSILAQKITPATLGALIAYYEHVTFTEGAIWNINSFDQWGVLELKG(V)
 308 375 8271.2075 (K)GAEAVDNHFTQTPLEDNIPLLGGLLSVWYNNFFGAQTHLVAPFDQYLHRFPAYLQQLSMESNGKSVTR(G)
 283 356 8941.5660 (R)YSVWSAIGLSVALYIGYDNFEAFLKGAEAVDNHFTQTPLEDNIPLLGGLLSVWYNNFFGAQTHLVAPFDQYLHR(F)
 271 356 10309.1741 (K)NMFGEFESWVGGRYSVWSAIGLSVALYIGYDNFEAFLKGAEAVDNHFTQTPLEDNIPLLGGLLSVWYNNFFGAQTHLVAPFDQYLHR(F)
 283 371 10939.5847 (R)YSVWSAIGLSVALYIGYDNFEAFLKGAEAVDNHFTQTPLEDNIPLLGGLLSVWYNNFFGAQTHLVAPFDQYLHRFPAYLQQLSMESNGK(S)

Fig. S8 MS digest results of glucose-6-phosphate isomerase (part 5). See descriptions on Page 64.

Table S2 Heat map of proteomics results in Table 4.1 (Page 67)

Gene	Description	No. of peptides	Average of B/A	Average of C/A	Average of D/A
Galactose metabolism					
GAL1	Galactokinase	9	0.870	0.587	3.783
GAL7	Galactose-1-phosphate uridylyltransferase	2	1.230	1.387	1.008
Glycolysis					
HXK1	Hexokinase-1	2	0.723	0.691	3.489
PGI1	Glucose-6-phosphate isomerase	3	1.032	1.330	9.891
PFK2	Phosphofructokinase	3	1.236	1.149	1.426
FBA1	Fructose-bisphosphate aldolase	8	1.094	0.871	2.304
TPI1	Triosephosphate isomerase	6	1.123	0.849	1.038
TDH	Glyceraldehyde 3-phosphate dehydrogenase	15	1.117	0.925	1.664
PGK1	Phosphoglycerate kinase	19	1.066	0.798	1.286
GPM1	Phosphoglycerate mutase 1	14	1.047	0.750	1.500
ENO	Enolase	19	1.185	0.912	2.176
PYK1	Pyruvate kinase	6	1.186	0.831	1.831
TCA cycle					
CIT1	Citrate synthase, mitochondrial	3	1.256	1.982	5.580

ACO1	Aconitase, mitochondrial	3	0.961	1.033	1.737
ATP synthesis					
ATP1	ATP synthase subunit alpha, mitochondrial	6	1.222	1.005	1.157
ATP2	ATP synthase subunit beta, mitochondrial	8	1.176	1.123	3.216
Amino-acid metabolism					
LEU1	3-isopropylmalate dehydratase	3	0.883	1.023	1.424
LEU2	3-isopropylmalate dehydrogenase	17	3.477	1.070	2.570
MET6	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	10	1.050	0.881	2.018
PDC	Pyruvate decarboxylase isozyme	12	1.300	1.118	1.894
Protein biosynthesis					
TIF	ATP-dependent RNA helicase eIF4A	3	1.408	0.752	1.655
TEF1	Elongation factor 1-alpha	8	0.910	0.758	1.507
RPL4	60s ribosomal protein L4	9	1.245	0.778	1.418
RPL19	60s ribosomal protein L19	2	1.218	1.114	2.995
Heat shock proteins					
HSP 12	12 kDa Heat shock protein	2	2.199	0.882	2.308
HSP 26	Heat shock protein 26	3	2.281	1.823	2.453
STI1	Heat shock protein STI1	2	1.363	0.485	3.450

Unknown

POR1	Mitochondrial outer membrane protein porin 1	4	1.033	0.808	2.785
SAM2	S-adenosylmethionine synthetase 2	2	1.271	0.624	2.125
YMR226C	Uncharacterized oxidoreductase YMR226C	2	1.856	1.260	3.051
SOD1	Superoxide dismutase [Cu-Zn]	2	6.360	3.942	7.910

*Error bar can refer to Table 4.1.

Color bar:

See description on Page 67.

Table S3 Data of growth curve Fig. 5.4 (Page 93)

WT-LHPA					
t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.345	0.307	0.334	0.329	0.020
8	1.219	1.129	1.257	1.202	0.066
12	4.070	3.760	3.870	3.900	0.157
16	6.640	6.070	6.530	6.413	0.302
20	8.280	7.530	8.170	7.993	0.405
24	9.020	8.270	8.820	8.703	0.388
28	9.130	8.470	9.380	8.993	0.470
32	9.210	8.720	9.360	9.097	0.335

<i>Δpxa1</i>- LHPA					
t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.200	0.200	0.200	0.409	0.018
8	0.388	0.416	0.422	1.215	0.058
12	1.158	1.214	1.274	3.807	0.325
16	3.440	3.920	4.060	7.050	0.425
20	6.560	7.310	7.280	8.503	0.193
24	8.280	8.620	8.610	9.207	0.133
28	9.060	9.240	9.320	9.397	0.203
32	9.170	9.460	9.560	9.503	0.304

<i>Δpxa2</i>- LHPA					
t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.332	0.368	0.344	0.348	0.018
8	1.215	1.458	1.224	1.299	0.138

12	4.060	4.520	4.010	4.197	0.281
16	6.590	7.410	6.380	6.793	0.544
20	8.070	8.620	8.190	8.293	0.289
24	8.950	9.550	8.760	9.087	0.412
28	8.930	9.740	9.120	9.263	0.424
32	9.220	9.730	9.260	9.403	0.284

Δpxa1&2- LHPA

t/h	1	2	3	Avg	SD
0		0.2		0.2	0
4	0.387	0.417	0.432	0.412	0.023
8	1.235	1.356	1.332	1.308	0.064
12	4.330	4.630	4.560	4.507	0.157
16	6.920	7.420	7.250	7.197	0.254
20	8.180	8.580	8.450	8.403	0.204
24	9.070	9.660	9.280	9.337	0.299
28	9.260	9.780	9.590	9.543	0.263
32	9.780	9.920	10.230	9.977	0.230

I have obtained permission from respective publisher to reproduce the work and figures in the thesis.

APPENDIX

List of publications

Li X, Chen W. 2014. Proteomics analysis of metabolically engineered yeast cells and medium-chained hydrocarbon biofuel precursors synthesis. *AMB Express* **4**:61.

Zhang J, Shi J, Lee BJ, Chen L, Tan KY, Tang X, Tan JY, **Li X**, Feng H and Chen WN 2013. Proteomic analysis of vascular smooth muscle cells with S- and R-enantiomers of atenolol by iTRAQ and LC-MS/MS. *Methods Mol Biol* **1000**: 45-52.