Research on the Stability of the Recombinant Production of Omega-3 Fatty Acids in Industrial Strains Using the Gene Cluster Isolated from

*Shewanella baltica* MAC1

by

Yi Fan

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Master of Science
in
Food Science

Guelph, Ontario, Canada

© Yi Fan, December 2015
Abstract

RESEARCH ON THE STABILITY OF THE RECOMBINANT PRODUCTION OF OMEGA-3 FATTY ACIDS IN INDUSTRIAL STRAINS USING THE GENE CLUSTER ISOLATED FROM *Shewanella baltica* MAC1

Yi Fan
University of Guelph, 2015
Advisor: Professor M. W. Griffiths

In past decades, there has been increasing awareness of the effects of omega-3 (n-3) fatty acids, especially EPA and DHA on human health. Realizing that these fatty acids are deeply involved in regulating human diseases as well as helping the proper development of the immune and nervous systems as well as visual acuity, more people began consuming omega-3 fatty acids in the form of nutritional supplements. So far, fish and fish oil are the main dietary sources of EPA and DHA. However, fish stocks have been depleted, therefore alternative sources of these fatty acids need to be explored. In this study, long-term stability of the plasmid carrying the 20 kb EPA/DHA biosynthesis gene cluster in the transformed *E. coli* EPI300T1 without antibiotic pressure as well as its recombinant production of EPA and DHA were studied. Moreover, molecular methods were used to transform *Streptococcus thermophilus* ST21 with plasmid pEDSB (modified plasmid vector pIL252 carrying the 20 kb insert), followed by a series of GC and GC/MS analyses to test recombinant production of EPA and DHA by *S. thermophilus* ST21. Our results showed recombinant production of EPA by *S. thermophilus*. However, production of EPA synthesized by transformed *S. thermophilus* was lower than that by recombinant...
*E. coli*. In addition, recombinant *E. coli* was able to stably maintain the 20 kb EPA/DHA gene cluster as compared to *S. thermophilus*. Our data confirmed transformation of the EPA/DHA gene cluster to gram-negative and gram-positive bacteria as well as its stability in *E. coli*.

**Key words:** EPA/DHA gene cluster, *Streptococcus thermophilus*, Gene cluster transformation, Gas Chromatography/Gas Chromatography-Mass Spectrometry
ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Dr. Mansel Griffiths, who gave me the opportunity to work in CRIFS in this project. In the past three years, he has always been supportive and patient. Thanks to all his instructive advice and valuable respect, I was able to learn and develop my own ideas so that I can finish the project independently.

I would like to thank my advisory committee members, Dr. Gisèle LaPointe and Dr. David Ma, who have shown their consistent encouragement and given out professional guidance since the project started.

I would like to thank my project supervisor, Dr. Amiri-Jami, who gave me considerable help by means of suggestion, comments and criticism. Her encouragement and unwavering support as a friend has sustained me through frustration and depression.

I would also like to thank all my colleagues, friends and classmates in CRIFS, thank you for helping me in the past amazing three years.

I can never be thankful enough for my family. My parents have always supported and encouraged me to follow my heart and pursue my dreams, even though they do not understand the things I am doing; even though they are now on the other side of this planet. I also thank my parents-in-law for supporting Justina and me, letting me feel loved and welcome as their own child.

I can never be thankful enough for my soul mate, my best friend, and my wife, Justina, who loves and supports me, cherishes and respects me unconditionally. I can never finish this thesis without her tolerance of my imperfection, especially my occasional grouchiness under stress. With her, all the tough times and frustration become easier to get through, all the joys and excitements become more meaningful. I also thank our lovely little princess, Penny. I am blessed to have you and your mom in my life. I really appreciate the support you brought me during writing the thesis, and all the joy and pride you brought to my life.

Lastly, I thank all the people in my life. Thank you for making my life meaningful.
1. Literature Review ............................................................................................................. 1

1.1 Introduction.................................................................................................................. 1

1.2 Polyunsaturated Fatty Acids (PUFAs): Eicosapentaenoic acid (EPA) and Docosahexaenoic acid (DHA)........................................................................................................... 5

1.2.1 Classification of PUFAs .......................................................................................... 5

1.2.2 Polyunsaturated Fatty Acid Synthesis in the Human body ...................................... 5

1.2.3 Dietary sources of n-3 and n-6 PUFAs .................................................................... 6

1.2.4 Documented EPA/DHA functions and Health Benefits ........................................ 10

1.3 EPA/DHA in marine bacteria ....................................................................................... 19

1.3.1 Possible Mechanism of PUFA biosynthesis in bacteria ......................................... 22

1.3.2 Polyketide synthase (PKS) pathway in marine bacteria ........................................ 24

1.3.3 The PKS pathway found in Shewanella baltica MAC1 .......................................... 26

1.3.4 Heterologous expression of the EPA/DHA biosynthesis gene cluster found in Shewanella baltica MAC1 .................................................................................................................. 30

1.4 Streptococcus thermophilus ......................................................................................... 32

1.4.1 A brief introduction of Streptococcus thermophilus ................................................. 32

1.4.2 Streptococcus thermophilus as probiotic .................................................................. 34

1.5 Prospects of genetically modified (GM) food ............................................................. 35

1.6 Research objectives ..................................................................................................... 38

2. Stability of EPA/DHA production by recombinant E. coli EPI300T1 no. 6 harboring plasmid pfBS-PS in consecutive cultivation .................................................. 40

2.1 Introduction .................................................................................................................. 40

2.2 Materials and methods ............................................................................................... 42

2.2.1 Bacterial strains and culture conditions .................................................................. 42

2.2.2 Test of the stability of plasmid pfBS-PS in E. coli EPI300T1 no.6 ............................ 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.3</td>
<td>Total lipid extraction of recombinant <em>E. coli</em> EPI300T1</td>
<td>46</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Gas chromatographic (GC) analysis of fatty acids</td>
<td>47</td>
</tr>
<tr>
<td>2.2.5</td>
<td>Statistical analysis</td>
<td>47</td>
</tr>
<tr>
<td>2.3</td>
<td>Results and discussion</td>
<td>48</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Stability of the recombinant plasmid pFB-S-PS in recombinant <em>E. coli</em> EPI300T1</td>
<td>48</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Recombinant bacterial EPA and/or DHA can be stably produced in recombinant EPI300T1 <em>E. coli</em> strains without selective pressure</td>
<td>50</td>
</tr>
<tr>
<td>2.4</td>
<td>General Discussion</td>
<td>55</td>
</tr>
<tr>
<td>3</td>
<td>The construction of recombinant <em>Streptococcus thermophilus</em> strain harboring the 20 kb EPA/DHA gene cluster</td>
<td>59</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>59</td>
</tr>
<tr>
<td>3.2</td>
<td>Material and methods</td>
<td>64</td>
</tr>
<tr>
<td>3.2.1</td>
<td>CRIFS stock culture confirmation</td>
<td>64</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Test of the transformability of <em>Streptococcus thermophilus</em> ST21</td>
<td>67</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Transformation of <em>Streptococcus thermophilus</em> ST21 with pIL252</td>
<td>70</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Transformation of the plasmid pIL252m + 20 kb EPA/DHA gene cluster in <em>Streptococcus thermophilus</em> ST21</td>
<td>73</td>
</tr>
<tr>
<td>3.2.5</td>
<td>Gas Chromatography (GC) and Gas Chromatography - Mass Spectrometry (GC-MS) analysis of the presumptive pEDSB transformed <em>Streptococcus thermophilus</em> clones</td>
<td>77</td>
</tr>
<tr>
<td>3.3</td>
<td>Results and discussion</td>
<td>78</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Confirmation of the identity of <em>Streptococcus thermophilus</em> ST21</td>
<td>78</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Antibiotic sensitivity and transformability of <em>Streptococcus thermophilus</em> ST21</td>
<td>80</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Vector plasmid pIL252 can be maintained and replicated in <em>Streptococcus thermophilus</em> ST21</td>
<td>84</td>
</tr>
</tbody>
</table>
3.3.4 Transformation of plasmid pEDSB (pIL252m + 20 kb EPA/DHA gene cluster) into *Streptococcus thermophilus* ST21 .................................................................................................................. 88

3.3.5 Gas Chromatography (GC) and Gas Chromatography - Mass Spectrometry (GC-MS) analysis of the pEDSB transformed *Streptococcus thermophilus* clones ........................................... 94

3.3.6 Stability of recombinant plasmid pEDSB in *Streptococcus thermophilus* STED20 .... 98

3.4 General discussion ........................................................................................................ 100

4 General discussion and future directions .................................................................... 109

4.1 Introduction ................................................................................................................ 109

4.2 General discussion .................................................................................................... 111

4.2.1 Plasmid stability .................................................................................................... 111

4.2.2 Recombinant production of EPA and DHA ....................................................... 117

4.3 Future directions ....................................................................................................... 119
List of Tables

Table 1.2.1.1 Average triacylglycerol fatty acid composition of various foods and oils (Erdman et al. 2012)……………………………………………………………………9

Table 2.2.1 Bacterial strains used in this study……………………………43

Table 2.2.2.2 Restriction enzyme digestion setup ………………………46

Table 2.3.2 The composition of EPA and DHA in total fat of the recombinant E. coli…54

Table 3.2.1.3A Primers used to amplify/determine housekeeping gene glcK and thrS…66

Table 3.2.1.3B1 PCR conditions used to amplify glcK in PCR……………66

Table 3.2.1.3B2 PCR conditions used to amplify thrS in PCR……………66

Table 3.2.2.5 Restriction enzyme digestion setup ………………………..70

Table 3.2.3.5 Restriction enzyme (BamHI) digestion setup ………………72

Table 3.2.4.4A Primers used to amplify gene pfaA………………………75

Table 3.2.4.4B PCR conditions used to amplify pfaA ……………………75

Table 3.2.4.5 Restriction enzyme (NotI HF) digestion setup…………………76

Table 3.3.2. Transformation Efficiency of pGKV 210 into S. thermophilus ST21……..83
List of Figures

Fig.1.1 The chemical structures and conversion between important polyunsaturated fatty acids. (Russell & Nichols, 1999)……………………………………………………………4

Fig.1.2.2 Essential fatty acids and their metabolites. ........................................ 6

Fig.1.2.3.3 Metabolites of n-3 fatty acids and n-6 fatty acids. ..........................16

Fig. 1.3.1 Phylogenetic tree based on the 16S rRNA gene sequence data of the PUFA-producing marine bacteria (Bowman et al. 1998) ................................. 21

Fig.1.3.2 “Anaerobic” fatty acids biosynthesis in Bacteria Steps of the pathway……24

Fig. 1.3.2.1 Putative EPA biosynthesis pathway by Shewanella polyketide synthase. (Wallis et al.2002)..................................................................................26

Fig. 1.3.3 Shewanella baltica MAC1 EPA/DHA gene cluster and its flanking region. (Amiri-Jami & Griffiths 2010) .................................................................29

Fig. 2.1 The structure of recombinant plasmid pfBS-PS (Amiri-Jami et al. 2014)....41

Fig. 2.3.1 A NotI-HF digestion result of the extracted plasmid from recombinant E. coli EPI300 T1 no.6 (day 0) .................................................................49

Fig. 2.3.1 B NotI-HF digestion result of the extracted plasmid (day 20) from recombinant E. coli. .................................................................49

Fig. 2.3.2.1 Gas chromatographs obtained from cells of recombinants.............51-52
Fig. 2.3.2.2 The yield of the recombinant EPA (mg of EPA/ g cell dry weight) by *E. coli* grown with chloramphenicol and without chloramphenicol………………53

**Fig. 3.1** Phylogenetic tree of *Lactobacillales* constructed on the basis of concatenated alignments of ribosomal proteins. ………………………………………...63

**Fig.3.3.1** Gel electrophoresis (1% agarose) of PCR amplification products of genes *glcK* and *thrS*. ………………………………………………………………………………………79

**Fig. 3.3.2** Gel electrophoresis (1% agarose) of plasmid pGKV 210 following restriction enzyme digestion………………………………………………………………………….82

**Fig. 3.3.3** Gel electrophoretogram (1% agarose) of plasmid pIL252 following restriction enzyme digestion. …………………………………………………………………………87

**Fig. 3.3.4.1** Gel electrophoretogram of amplicons following colony PCR targeting the *pfaA* gene. …………………………………………………………………………………89

**Fig. 3.3.4.2a.** Electrophoretograms of plasmids extracted from *Lactococcus lactis* subsp. *cremoris* MG1363 and digested with *NotI* HF. …………………………………91

**Fig. 3.3.4.2b.** Electrophoretograms obtained following restriction enzyme (*NotI* HF) digestion of plasmids extracted from pIL252m + 20kb inserted mutants………………92

**Fig.3.3.5.1** GC results……………………………………………………………………96

**Fig.3.3.5.2** GC-MS results……………………………………………………………… 97

**Fig. 3.3.6** Gel electrophoretogram following restriction enzyme *NotI* HF digestion of plasmids extracted from the re-subcultured *S. thermophilus* STED20. ………………99
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>Arachidonic acid</td>
</tr>
<tr>
<td>ACC</td>
<td>Acetyl coenzyme A carboxylase</td>
</tr>
<tr>
<td>ACP</td>
<td>Acyl carrier protein</td>
</tr>
<tr>
<td>ALA</td>
<td>Alpha-Linolenic acid</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine tri phosphate</td>
</tr>
<tr>
<td>AT</td>
<td>Acyl-transferase</td>
</tr>
<tr>
<td>bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>C</td>
<td>Carbon</td>
</tr>
<tr>
<td>Cas9</td>
<td>CRISPR associated protein 9</td>
</tr>
<tr>
<td>CLF</td>
<td>Chain Length Factor</td>
</tr>
<tr>
<td>cm</td>
<td>Centimetre</td>
</tr>
<tr>
<td>CoA</td>
<td>Coenzyme A</td>
</tr>
<tr>
<td>CVD</td>
<td>Cardiovascular disease</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>DHA</td>
<td>Docosahexaenoic acid</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
</tbody>
</table>
dNTP 2′ – deoxyribonucleoside 5” - triphosphate

*E. coli* *Escherichia coli*

EB Elution buffer

EDTA Ethylenediaminetetraacetic Acid

EFA Essential fatty acids

EPA Eicosapentaenoic acid

ETB Ethidium Bromide

FA Fatty acid

*Fab* Fatty acid biosynthesis

FAS Fatty acid biosynthesis

FID Fume ionization detector

FOS Fosmid

*g* Gravitational acceleration

GC Gas chromatograph

GC-MS Gas chromatographic – Mass spectrometry

*glcK* *Glucose kinase*
GM17 supplemented with Glucose

h Hour

H₂ Hydrogen

HCl Hydrogen Chloride

HD Hydroxydecarbonyl Dehydratase

HDL-C High-density lipoproteins-cholesterol

HMW High molecular weight

KAS 3-oxoacyl carrier protein synthase

kb Kilobase

KC1 Potassium Chloride

KR Ketoacyl Reductase

KS Ketoacyl-ACP synthase

L Liter

L. Lactococcus

LA Linoleic acid

LBA LB Agar

LDL-C Low density lipoproteins-cholesterol
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMP</td>
<td>Low melting point</td>
</tr>
<tr>
<td>LPUFA</td>
<td>Long chain polyunsaturated fatty acid</td>
</tr>
<tr>
<td>LTB</td>
<td>Leukotrienes</td>
</tr>
<tr>
<td>m</td>
<td>Metre</td>
</tr>
<tr>
<td>M</td>
<td>Mole</td>
</tr>
<tr>
<td>MA</td>
<td>Marine agar</td>
</tr>
<tr>
<td>MB</td>
<td>Marine broth</td>
</tr>
<tr>
<td>MCT</td>
<td>Malonyl CoA-ACP Transferase</td>
</tr>
<tr>
<td>mg</td>
<td>Milligram</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium Chloride</td>
</tr>
<tr>
<td>MgSO₄</td>
<td>Magnesium Sulfate</td>
</tr>
<tr>
<td>Min</td>
<td>Minute</td>
</tr>
<tr>
<td>ml</td>
<td>Milliliter</td>
</tr>
<tr>
<td>mm</td>
<td>Millimetre</td>
</tr>
<tr>
<td>ms</td>
<td>Millisecond</td>
</tr>
<tr>
<td>mM</td>
<td>Millimole</td>
</tr>
<tr>
<td>MUFA</td>
<td>Monounsaturated Fatty Acid</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>nm</td>
<td>Nano meter</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>OL</td>
<td>Oleic acid</td>
</tr>
<tr>
<td>ORFs</td>
<td>Open reading frames</td>
</tr>
<tr>
<td>Par</td>
<td>Partitioning</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PG</td>
<td>Prostaglandin</td>
</tr>
<tr>
<td>Pfa</td>
<td>Polyunsaturated fatty acid</td>
</tr>
<tr>
<td>PKS</td>
<td>Polyketide Synthase</td>
</tr>
<tr>
<td>PPtase</td>
<td>Phosphopantetheinyl Transferase</td>
</tr>
<tr>
<td>PUFA</td>
<td>Polyunsaturated Fatty Acid</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal Ribonucleic acid</td>
</tr>
<tr>
<td>S</td>
<td><em>Streptococcus</em></td>
</tr>
<tr>
<td>Sh</td>
<td><em>Shewanella</em></td>
</tr>
<tr>
<td>Sec</td>
<td>Second</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris/acetate buffer</td>
</tr>
<tr>
<td>Tn5</td>
<td>Transposon five</td>
</tr>
<tr>
<td>thrS</td>
<td>Threonine</td>
</tr>
<tr>
<td>Tris-HCl</td>
<td>Tris-hydrochloride</td>
</tr>
<tr>
<td>TXs</td>
<td>Thromboxanes</td>
</tr>
<tr>
<td>U</td>
<td>Unite</td>
</tr>
<tr>
<td>µg</td>
<td>Microgram</td>
</tr>
<tr>
<td>µl</td>
<td>Microliter</td>
</tr>
<tr>
<td>µM</td>
<td>MicroMole</td>
</tr>
<tr>
<td>V</td>
<td>Volt</td>
</tr>
</tbody>
</table>
1. Literature Review

1.1 Introduction

Polyunsaturated fatty acids (PUFAs) are fatty acids that contain more than one double bond in their backbone chain. According to their chemical structures, PUFAs can be classified in various groups, among which the omega-3 fatty acids and the omega-6 fatty acids are the most important. For example, linoleic acid (LA, C18:2n-6) and arachidonic acid (ARA, C20:4n-6) are the most commonly discussed compounds when we talk about the omega-6 fatty acids; whereas α-linolenic acid (ALA, C18:3n-3), eicosapentaenoic acid (EPA, C20:5n-3) and docosahexaenoic acid (DHA, C22:6n-3) have attracted most attention in the omega-3 family. Essential fatty acids (EFAs) are the fatty acids that humans cannot synthesize by themselves, but have to obtain from food to maintain body wellness. Among all PUFAs, only α-linolenic and linoleic acids are still considered as “essential fatty acids” by most nutritionists. However, because of the limited efficiency of biosynthesis in the human body, arachidonic acid (converted from linoleic acid), EPA and DHA (converted from α-linolenic acid) can also be considered as “conditional essential” (structures shown in Fig. 1.1).

Research on the role of PUFAs in the human body started from the late 1920s and the early 1930s, when scientists found that two vitamin-like substances could be detected in dietary fats (Alfin-Slater & Aftergood 1968). They were soon identified as linoleate (LA) and α-linolenate (ALA) (Alfin-Slater & Aftergood 1968). However, it was not until research on prostaglandins (PG), leukotrienes (LTB0) and thromboxanes (TXB) started fifty years later that PUFAs gained more attention (Simopoulos 1989). After 1980, when
the beneficial association between PUFA-rich diets and human disease (e.g., cardiovascular disease, brain function disorder, retina damage, and cancers) was made did researchers become aware of the great potential of PUFAs (Siddiqui et al. 2004; Fontani et al. 2005). Considering all the benefits that PUFAs have to the human body, the wide application of PUFAs, especially EPA and DHA, in the pharmaceutical, food and feed industries in recent years is not surprising.

It is interesting that although several PUFA-producing bacterial strains were reported in the 1970s (Oliver & Colwell 1973), not much attention was placed on PUFAs of bacterial origin. Fortunately, by analyzing the DHA and EPA found in marine bacteria, the potential of these microbial PUFAs was realized, which resulted in the initiation of research into PUFA-producing bacteria (Delong & Yayanos 1986).

Nevertheless, these bacteria were still not given enough attention until the 1990s. The breakthrough occurred in 1996 when a Japanese scientist, Kazunaga Yazawa, first successfully cloned the 5-gene cluster involved in the biosynthesis of EPA from a *Shewanella* sp., (Yazawa 1996), subsequently identified as *Shewanella pneumatophori* SCRC-2738 (Hirota 2005). Meanwhile, with the development of molecular techniques and bioinformatics, research on PUFA-related genes was stimulated and led to the isolation and cloning of the DHA-synthesis gene cluster from *Moritella marina* strain MP-I (Morita et al. 2000), and a new EPA-synthesis related 4-gene cluster from *Photobacterium profundum* SS9 (Allen & Bartlett 2002). A more efficient PUFA-gene cluster was found recently in *Shewanella* (Amiri-Jami et al. 2006), which when cloned into *E. coli* resulted in a strain that was able to efficiently produce both EPA and DHA (Amiri-Jami & Griffiths 2010). It is not hard to understand why attention has been
focused on bacterial production of EPA and DHA. First and foremost, cultivation of some bacteria is simple and fast. Moreover, compared to eukaryotic organisms (e.g. yeast, Arabidopsis thaliana, zebrafish, laboratory mice), genetic manipulation of prokaryotic bacteria, especially in a well-characterized organism like E. coli, is generally straightforward making it possible to incorporate the 35 kb EPA-synthesis gene cluster from Shewanella, which opens the possibility of the future use of the EPA and DHA gene cluster for commercial bacterial production of omega-3 fatty acids. The EPA- and DHA-gene cluster cloned into a relevant microorganism has great potential to increase use of EPA and DHA. Considering the limitation of the current source of n-3 fatty acids and n-6 fatty acids (animals and plants), microbial production may be an ideal alternative source in the future. So far, microalgae, such as Phaeodactylum, Porphyridium, Nannochloropsis salina, Isochrysis, Chlamydomonas have been commercially used to produce PUFAs, especially EPA and DHA (Li et al. 2002). Nevertheless, compared to traditional dietary omega-3 sources (cold water marine fish and fish oil), the industrial mass production of microalgae EPA and DHA is costly due to the technological bottleneck associated with fat filtration, separation, and refining (Klok et al. 2014). To meet the huge and increasing demand of the EPA and DHA market, a novel output-stable, edible, highly efficient and low-cost alternative source is urgently needed. Thus, in the near future, recombinant probiotics capable of producing omega-3 fatty acids may be applied in the food and feed industry. In this case, a new dietary habit may be established to help human beings build healthier lives.
Fig. 1.1 The chemical structures and conversion between important polyunsaturated fatty acids. The conformation of the acyl chains from α-linolenic acid to docosahexaenoic acid (DHA) (n-3 family); and from linoleic acid to arachidonic acid (n-6 family) are presented in a stylized format (from Russell & Nichols, 1999).
1.2 Polyunsaturated Fatty Acids (PUFAs): Eicosapentaenoic acid (EPA) and Docosahexaenoic acid (DHA)

1.2.1 Classification of PUFAs

PUFAs can be classified according to the position of the first double bond from the end of their carbon chain. In this way, PUFAs can be classified as n-3, n-6, n-7 and n-9, etc.

Also, PUFAs have been classified as 1) the Methylene-Interrupted Polyenes and 2) the Conjugated Fatty Acids. For the Methylene-Interrupted Polyenes, there are 2 or more cis double bonds that are separated from each other by a single methylene in this group, and, as its name suggests, the Conjugated Fatty Acids family, contain double bonds that are adjacent to each other.

1.2.2 Polyunsaturated Fatty Acid Synthesis in the Human body

In humans, fatty acids are synthesized de novo from acetyl coenzyme A (CoA) by a set of fatty acid synthetases. Units of malonyl CoA, a metabolite of acetyl CoA, are added unit by unit to elongate the fatty acid chain to form palmitic acid (C16:0). From then on, depending on our body’s dietary needs, fatty acid elongation and desaturation is performed by an elongase and a set of desaturase enzymes. However, humans do not have the enzyme to desaturate sites below the n-7 carbon. Therefore, n-3 and n-6 fatty acids are essential to human beings. Being modified by a set of desaturases (D5, D6 and D9 desaturase) and an elongase, α-linolenic acid (ALA, C18:3n-3) can become eicosapentaenoic acid (EPA, C20:5n-3) or docosahexaenoic acid (DHA, C22:6n-3); while linoleic acid (LA, C18:2n-6) can convert to arachidonic acid (ARA, C20:4n–6), as shown in Fig.1.2.2.
1.2.3 Dietary sources of n-3 and n-6 PUFAs

Lipid (fat) is an important dietary component; not only as a critical source of metabolic energy and carrier of other nutrients such as fat-soluble vitamins, but also fats, especially essential fatty acids, serve as precursors of other bioactive compounds. Fatty acids also serve as the most critical component of cell membranes. Known as a dynamic structure, the cell membrane literally supports almost all the pivotal life processes (e.g. cell signaling, nutrient transportation, endocytosis, and exocytosis) (Alberts et al. 2007).
As mentioned, due to the lack of an enzyme able to synthesize a double bond in the n-3 and n-6 positions of the carbon chain of a fatty acid, humans have to obtain the essential fatty acids α-linolenic acid and linoleic acid from food to prevent deficiency. However, due to the low rate of conversion from α-linolenic acid to EPA (6%) and DHA (3.8%), and from linoleic acid to arachidonic acid (5%), n-3 fatty acids should be considered “conditionally essential fatty acids” (Cunnane 2003; Gerster 1998). Therefore, to maintain optimal health, USDA recommends an intake of 1.6 g/day and 1.1 g/day of n-3 fatty acids for men and women, respectively (Trumbo et al. 2001).

Among all the dietary sources of fat, vegetables and seafood are still the main human sources of n-3 and n-6 PUFAs. The fatty acid profile of vegetable oils can vary widely (Table 1.2.2.1). Safflower, sunflower and corn oils are very good sources of linoleic acid but cannot be considered as a significant α-linolenic acid source. By contrast, flaxseed, linseed and canola oils have a high percentage of α-linolenic acid but are relatively low in linoleic acid. For example, in flax seed oil, the content of α-linolenic acid can reach as high as 57% with only 13% linoleic acid, and in sunflower oil, the ratio between linoleic acid and α-linolenic acid can reach as high as 70:1 (Simopoulos 1994).

However, marine-derived n-3 fatty acids, such as marine fish and algae oil, are still the main sources of our dietary EPA and DHA. The content of these n-3 fatty acids also vary from different fish organs. For example, in cod liver oil, vitamin A and vitamin D are the main ingredients (Simopoulos 1994), whereas, in fish body lipids, between 25% to 59% of the fat is composed of n-3 PUFAs. Notably, marine bacteria and marine algae are very good sources of n-3 fatty acids. Algae such as *Phaeodactylum* spp., *Porphyridium* spp., *Nannochloropsis salina*, *Isochrysis* spp., *Chlamydomonas* spp. have been used to produce...
PUFAs (Li et al. 2002). Limited by experimental design and analytical methods, it was argued that bacteria cannot synthesize PUFAs by themselves (Scheuerbrandt & Bloch 1962). However, with the development of technology, it was realized that bacteria, such as *Shewanella* spp., and *Colwellia psychrerythraea*, from deep sea environments where they are subjected to low temperatures and high pressure, have developed pathways to synthesize long-chain polyunsaturated fatty acids to allow them to survive in these extreme environments by increasing membrane lipid fluidity (Nichols et al. 1995). By successfully cloning the gene cluster responsible for EPA/DHA synthesis from these marine bacteria, bacterial recombinant n-3 fatty acids may become a new and more environmental friendly dietary source of these fats (Yazawa 1996; Morita et al. 2000; Allen & Bartlett 2002; Amiri-Jami et al. 2006; Amiri-Jami et al. 2014).
Table 1.2.1.1 Average triacylglycerol fatty acid composition of various foods and oils

(Erdman et al. 2012)

<table>
<thead>
<tr>
<th>FOOD/OIL</th>
<th>Average Fat %</th>
<th>Saturated/monounsaturated</th>
<th>Essential fatty acid</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Palmitic acid C16:0</td>
<td>Stearic acid C18:0</td>
</tr>
<tr>
<td>Almond oil</td>
<td>100</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Avocado oil</td>
<td>100</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Beef tallow</td>
<td>100</td>
<td>29</td>
<td>20</td>
</tr>
<tr>
<td>Butter</td>
<td>81</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>Canola</td>
<td>100</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Cashew nut</td>
<td>68</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>Coconut oil</td>
<td>100</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>Corn oil</td>
<td>100</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Cottonseed oil</td>
<td>100</td>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>Flaxseed oil</td>
<td>100</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Grapeseed oil</td>
<td>100</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Hazelnut oil</td>
<td>100</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Hemp oil</td>
<td>100</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Herring (menhaden)*</td>
<td>16-25</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>Mackerel</td>
<td>25</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>Milk (cow’s)</td>
<td>3.5</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>Olive oil</td>
<td>100</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>Palm Oil</td>
<td>100</td>
<td>45</td>
<td>5</td>
</tr>
<tr>
<td>Palm kernel oil</td>
<td>100</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Pork fat (lard)</td>
<td>100</td>
<td>28</td>
<td>13</td>
</tr>
<tr>
<td>Safflower seed oil</td>
<td>100</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Salmon</td>
<td>13</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>Sesame oil</td>
<td>100</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Soybean oil</td>
<td>100</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Sunflower seed oil</td>
<td>100</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Walnut</td>
<td>63</td>
<td>7</td>
<td>2</td>
</tr>
</tbody>
</table>

Note: The percentage given are approximations, climates, species, and other environmental factors cause great variations.

Trace: less than 1% detected; -, not detectable.

* Menhaden herring oil contains 11% EPA and 9% DHA, but Norwegian herring oil has 13% C20:1n-9, 21% C22:1n-11, 7% DHA. Depending on fishing location, mackerel oil is similar to menhaden or Norwegian herring.
1.2.4 Documented EPA/DHA functions and Health Benefits

“We are what we eat”, said Anthelme Brillat Savarin; indicating how our food significantly affects our health. It is known that PUFAs, especially n-3 fatty acids and n-6 fatty acids, are inextricably involved in human lipid-related metabolism. Many human diseases are linked, either directly or indirectly, to people’s dietary intake of PUFAs.

1.2.4.1 Essential fatty acid deficiency and n-3 fatty acid deficiency

The lack of essential fatty acids in the diet can lead to deficiency. The symptoms related to this deficiency were documented in the 1950s when infants were inadvertently given an essential-fat-free diet by their parents. Serious skin problems, including dry, leathery skin, desquamation and perianal irritation developed, and not until linoleic acid was added to the feed were these symptoms alleviated (Hansen et al. 1958). Later on, more signs of essential fatty acid deficiency including dry scaly rash, decreased growth in infants and children, increased susceptibility to infection, and poor wound healing were documented (Jeppesen et al. 1998). Unlike linoleic acid deficiency, very few α-linolenic acid deficiencies in humans were documented, and most evidence of the nutritional importance of this fatty acid was obtained from animal studies (Anderson & Connor). Even in the documented human cases (Meng 1983; Stein et al. 1983; Bjerve et al. 1987), symptoms may also have been caused by other nutrient deficiencies (e.g. Vitamin E, DHA, for example). To avoid deficiency symptoms, researchers estimated that the minimal daily requirement of linoleic acid in adults should be ~0.2-0.3% of energy (Meng 1983; Stein et al. 1983; Bjerve et al. 1987). In 2008, the recommended intake suggested by the World Health Organization of omega-6 fatty acid and omega-3 fatty acid were 5-8% of energy and 1-2% of energy, respectively.
The biological benefits of linoleic acid to α-linolenic acid also depend on their conversion to n-3 long chain polyunsaturated fatty acids (LCPUFAs). It has been shown that the synthesis of LCPUFAs is regulated by a D6 desaturase enzyme, and all three classes of the C18 unsaturated fatty acids (C18:3n-3, C18:2n-6, C18:1n-9) can serve as a substrate for the enzyme. However, as the D6 desaturase prefers more highly unsaturated fatty acids, the preferential desaturation order will be: C18:3n-3 > C18:2n-6 > C18:1n-9. Therefore, a high level of n-6 fatty acid consumption can suppress the desaturation and elongation of α-linolenic acid to EPA and DHA and thus lead to n-3 fatty acid deficiencies. Based on the essential fatty acid metabolism and the competitive desaturation of n-3, n-6, and n-9 fatty acids by D6 desaturase, it is important to keep a proper ratio of linoleic acid to α-linolenic acid in the diet. Arguing that the ratio of linoleic acid to α-linolenic acid in the Western diet is too high, scientists recommended a ratio from 2:1 to 5:1 to avoid n-3 fatty acid deficiency (Kris-Etherton et al. 2000; Simopoulos 2008; Holman 1998). Since the fat content of seed oils is high in n-6 fatty acid, to achieve the recommended ratio, more marine oils that contain a greater amount of n-3 fatty acids should be included in our diet.

Fatty acids that are not considered as essential for adults might be essential to infants, especially preterm babies who did not receive a sufficient intrauterine supply of arachidonic acid and DHA (Lapillonne & Jensen 2009). Scientists assume that arachidonic acid and DHA in breast milk contribute to an infant’s (both term and preterm) early development (Agostoni et al. 2009). Without adding arachidonic acid and DHA into infant formula, there might not be enough converted essential fatty acids, especially DHA, to meet the requirement for proper brain and visual development (Cunnane et al. 2000).
1.2.4.2 EPA/DHA and cardiovascular diseases (CVD)

In the 1970s, scientists found that coronary occlusion is uncommon in the Greenland Eskimo population. Compared to the average Danish Eskimo, the Greenland Eskimo diet contains more protein, less carbohydrate, and a significantly higher level of marine-derived polyunsaturated fatty acids, especially EPA. They assumed that the LCPUFAs in marine mammals might help lower blood cholesterol and $\beta$-lipoprotein, and thus lower their morbidity from coronary atherosclerotic diseases (Bang et al. 1976). This finding led to an explosion of fish oil-CVD research. The first clinical cardiovascular trial to study the benefit of fish oils was reported in 1989 (Burr et al. 1989). From then on, more and more clinical trials proved that a diet rich in fish oil can prevent sudden cardiac death in humans, and also significantly help reduce the mortality rate of cardiovascular diseases (de Lorgeril et al. 1994; GISSI-Prevenzione trial centres. 1999). More recently, according to a meta-analysis of 11 prospective cohort studies (involving 222,364 people, followed on average for 11.8 years), the risk ratio for CVD mortality for people who consumed fish once per week, 2-4 times a week, or more than 5 times a week was 15%, 23%, and 38% lower, respectively, than for people who never ate fish or ate fish less than once per month (He et al. 2004). An infarct survival trial in India showed that compared to the placebo group, an EPA/DHA-rich diet (1.8g/d) decreased total cardiac events from 34.7% to 24.5% ($P<0.01$) and death due to heart attacks from 22.0% to 11.4% ($P<0.05$) (Singh et al. 1997). Furthermore, a 17-year-long study supported the fact that high blood levels of DHA were strongly associated with a reduced risk of sudden cardiac death (Albert et al. 2002).

How EPA and DHA reduce cardiovascular disease risk is still being studied. So far,
scientists have proposed ways that fish oils can help (i) prevent arrhythmias, (Connor 2000) (ii) alter the serum lipid profile by decreasing triglyceride and modestly improve the low-density lipoprotein cholesterol (LDL-C) and high-density lipoprotein cholesterol (HDL-C) concentrations, (Balk et al. 2006), (iii) retard patients’ endothelial dysfunction, (Schiano et al. 2008), (iv) slightly lower blood pressure and platelet aggregation, (Geleijnse et al. 2002; Hornstra 2001) or (v) alleviate inflammation (Calder 2001).

1.2.4.3 EPA/DHA and inflammation

Inflammation is part of the body’s immediate response to infection or injury. It prompts the immune system to eliminate foreign pathogens or toxins and to repair damaged cells. Normally it can cause an increase in blood flow and permeability across blood capillaries to allow more bioactive molecules (e.g. antibodies and cytokines) to reach the infected/damaged tissue. However, the inflammatory response should be controlled, or it can cause acute/chronic human disease (e.g. rheumatoid arthritis, systemic lupus erythematosus, allergic rhinitis and even cancer) (Wang et al. 2010). It has been reported widely that n-3 fatty acids and n-6 fatty acids are heavily involved in regulating inflammatory response. To initiate inflammation in the human body, we need the release of pro-inflammatory cytokines (e.g. IL-1, TNF-α), which leads to the activation of phospholipase A2 (PLA2); resulting in the breakdown of cell membrane phospholipids to arachidonic acids. Later on, following association with cyclooxygenase (COX) and lipoxygenase (LOX), arachidonic acids are converted into eicosanoids, which are mediators of inflammation (e.g. 2-series prostaglandin and 4-series leukotrienes), and thus exacerbate an inflammation response (Boilard et al, 2010).

As mentioned before, n-3 fatty acids compete for the active site of the D6 desaturase with
n-6 fatty acids. Thus, exposure to n-3 fatty acids results in changing the cell membrane lipid profile, which results in an increased content of EPA and DHA. Therefore, by replacing the n-6 fatty acids, EPA/DHA can help relieve inflammation by decreasing the production of arachidonic-derived pro-inflammatory factors like PGE$_2$ (Healy et al. 2000), TAB$_2$ (Caughey et al. 1996), and LTB$_4$ (Sperling et al. 1993) in inflamed cells.

As a substrate of COX and LOX, EPA also serves as a precursor for eicosanoids that are different from those derived from arachidonic acid (Fig. 1.2.3.3). Compared to metabolites of arachidonic acid (e.g. 2-series prostanoids and 4-series leukotrienes), mediators formed from EPA (e.g. 3-series prostanoids and 5-series leukotrienes) are believed to be less potent mediators of inflammation. For instance, PGE$_3$ (derived from EPA) serves as a less potent inducer of both IL-6 production in macrophages to stimulate inflammation than PGE$_2$ (derived from arachidonic acid) (Bagga et al. 2003); and when inducing the adhesion and activation of leukocytes on the endothelium, LTB$_3$ (derived from EPA) is 10- to 100-fold less active than LTB$_4$ (derived from arachidonic acid) (Hawkes et al. 1991). Recently, it has also been reported that resolvins synthesized from EPA (E-series resolvins) and DHA (D-series resolvins) protect sites of inflammation from neutrophils (Serhan et al. 2002), whereas protectins (derived from DHA) were also found to inhibit the production of pro-inflammatory cytokines, TNF-$\alpha$ and IL-1$\beta$ (Adkins and Kelley 2010).

Fish oils have been shown to alleviate the inflammatory response by inhibiting the activation of NF-$\kappa$B (nuclear factor kappa-light-chain-enhancer of activated B cells), which induces a series of inflammation-related genes (Lo et al. 1999). Besides their
effects on production of TNF-α and IL-1 mentioned above, pure EPA and DHA can also inhibit the production of a series of inflammatory proteins including COX-2, IL-6, IL-8, and IL-12 in different human cell lines (Babcock et al. 2002; Draper et al. 2011; Lee et al. 2001).
Fig.1.2.3.3 Metabolites of n-3 fatty acids and n-6 fatty acids. pg: prostaglandin; pgi: prostacyclin; tx: thromboxane; lt: leukotriene. Orange sites indicate pro-inflammatory factors, whereas green sites indicate anti-inflammatory factors.

1.2.4.4 EPA/DHA and fetal development

A fish-oil rich diet or fish oil supplementation during pregnancy can benefit both the baby and mother. As mentioned above, the lack of essential fatty acids in infant formula leads to an essential fatty acid deficiency. However, EPA/DHA supplementation is important during pregnancy as well. During pregnancy, EPA and DHA are transferred to the fetus at rates dependent on the amount ingested by the mother (Helland et al. 2008).
Several studies proved that an adequate maternal dietary fish oil supplementation helps babies develop brain and retina functions, as DHA plays a key role in early brain and retina development. As early as the third trimester, a particularly high amount of DHA accumulates in fetal tissues, especially in the brain and retina (Clandinin et al. 1980). It is reported that children with mothers who had taken DHA supplementation during pregnancy have better problem-solving skills by the age of 9 months (Judge et al. 2007).

It is also well known that preterm labor is associated with many infant diseases (e.g., chronic lung disease) and even neonatal death. It is believed that EPA and DHA supplementation during pregnancy can help alleviate inflammation within the uterus to prevent preterm labor, and thus leads to an optimal pregnancy length (Olsen et al. 2008; Judge et al. 2007). Furthermore, clinical trials have shown that intake of EPA and DHA during pregnancy can also slightly increase infant birth weight and decrease the risk of delivering a low birth weight baby (Szajewska et al. 2006).

Fish oils can also help the mothers. As mentioned before, during pregnancy, mothers transfer DHA to their babies to support their early neurological development. However, without adequate fish oil intake, the lack of DHA may increase the risk of maternal postpartum depression (PPD). A multinational research study indicated that the concentrations of DHA in mother’s milk and seafood consumption were inversely correlated to the rate of postpartum depression (Hibbeln, 2002).

1.2.4.5 EPA/DHA and the visual system

Omega-3 long-chain polyunsaturated fatty acids, especially DHA, have a positive effect on neural and retinal development. Fetal retinal development can start as early as 6.5
weeks of gestational age (Hollenberg & Spira 1973). Therefore, adequate maternal fish oil supplementation is vital to fetal retina development. As a major component of membrane phospholipids in retinal photoreceptors, especially in photoreceptor outer segments, the content of DHA can reach as high as 65%, which makes it the highest among all neural subcellular components. The exact role of DHA in the development and function of the retina remains unclear, however, it is believed that a high DHA lipid profile can increase the degree of membrane fluidity, and thus help retinal cells function optimally. A low level of DHA in the photoreceptor system was always associated with impaired electroretinogram responses and decreased visual acuity (Burdge 1998).

Besides retina development and function requirements, it is indicated by epidemiology and clinical research that the n-3 LCPUFAs contribute to cytoprotective and cytotherapeutic actions by promoting the anti-angiogenic and neuroprotective mechanism within the retina. Effects of environmental exposures (such as oxidative stress, inflammation, cellular signaling mechanisms and aging) that cause vasoproliferative and neurodegenerative retinal diseases may be attenuated by the n-3 LCPUFAs. During the modulating metabolic processes, the production of molecules like eicosanoids, angiogenic factors, cyclic nucleotides, neurotransmitters and neuromodulators, pro-inflammatory and immunoregulatory cytokines, and inflammatory phospholipids are all regulated by the n-3 LCPUFAs (SanGiovanni & Chew, 2005).

1.2.4.6 EPA/DHA and brain

The last trimester of fetal life and the first two years after birth is a critical time for early brain development. During the “brain growth spurt”, n-3 LCPUFAs, especially DHA, accumulate in brain tissues, indicating that they might play a significant role in the early
development of certain neural functions (e.g. sensory, perceptual, cognitive, and motor neural system) (Martinez 1992; McCann & Ames 2005).

There is now a considerable literature pointing out that PUFAs might have great beneficial effects on various affective disorder diseases, such as epileptic seizures, depression or bipolar and other behavioral diseases (Lauritzen et al. 2000). Also, PUFA-rich diets may even help prevent or attenuate the negative effects of Alzheimer’s disease, which is, at present, incurable. The loss of DHA in the brains of patients with Alzheimer’s disease is accompanied by a decrease in memory and cognitive ability (Söderberg et al. 1991). There is also a negative correlation between the ratio of n-3 and n-6 fatty acids in erythrocyte membranes and cognitive decline among the elderly (Heude et al. 2003). Also, animal studies indicate that memory damage in normal aging is associated with decreased DHA concentration in cerebral cells (Delion et al. 1997; Favrelère et al. 2000). DHA may also be involved in S-adenosylhomocysteine (SAM) metabolism in Alzheimer’s disease; offering a possible mechanism of how PUFAs contribute to brain function (Selley 2007). In addition, PUFAs are also positive factors in the prevention and treatment of Parkinson’s disease (Tanriover et al. 2010) and cognitive function in pre-school children and the elderly (Dalton et al. 2009; Chiu et al. 2012).

1.3 EPA/DNA in marine bacteria

For decades it has been widely accepted that (unlike plants and animals) bacteria did not contain polyunsaturated fatty acids (PUFAs) as acyl functions of the complex lipids in their membranes (Harwood & Russell 1984). So far, the most common source of PUFAs is marine fish. Although several bacterial strains were reported to produce polyunsaturated fatty acids (Oliver & Colwell 1973), limited by the analytical technology,
the result did not receive much attention. Later on, eicosapentaenoic acid (20:5n3; EPA) was found in the marine bacterium *Flexibacter polymorphus*, and confirmed by gas chromatography–mass spectrophotometry (Perry et al. 1979). With such convincing methodology, the result challenged the assumption that bacteria do not harbor both aerobic and anaerobic pathways to synthesise PUFAs at the same time, and in the 1990s, it was proved conclusively that both aerobic and anaerobic pathways of fatty acid biosynthesis could operate within a single species (Wada et al. 1989). With more and more evidence, it is now widely accepted that some bacterial species, especially marine species, have the ability to produce polyunsaturated fatty acids. For example, *Shewanella gelidimarina* can produce ARA (Nichols & Russell 1996); *Shewanella hanedai* produces EPA; and *Colwellia psychrerythraea* yields both EPA and DHA (Bowman et al. 1998).

Recently, with the development of isolation and analytical methodologies, more and more EPA- and DHA-producing bacteria have been discovered in deep marine environments. Taxonomically speaking, PUFA-producing bacteria now can be classified by the phylogenetic investigations of 5S and 16S rRNA sequences. A close relationship has been established between the genera of common Gram-negative marine bacteria: *Colwellia, Shewanella, Alteromonas, Pseudoalteromonas* and *Ferrimonas*. Among them, *Colwellia* and *Shewanella* are now considered as the main agents for producing EPA/DHA in the ocean (Russell & Nichols 1999).
Fig. 1.3.1 Phylogenetic tree based on the 16S rRNA gene sequence data of the PUFA-producing marine bacteria (taken from Bowman et al. 1998).
1.3.1 Possible Mechanism of PUFA biosynthesis in bacteria

To adapt to low temperature and high pressure, the fatty acid synthesis system is crucial to marine bacteria. At low temperatures, the high melting point (saturated and straight chain) lipids in the cell membrane will crystallize. This crystallization decreases the fluidity of the cell membrane, and thus leads to disruption of the cell membrane’s physiological functions (Quinn 1985). Compared to saturated and straight chain fatty acids, unsaturated and branched chain fatty acids have lower melting points. To increase membrane fluidity, marine bacteria have developed strategies to increase the content of the low melting point (unsaturated and branched chain) fat in their cell membranes to maintain normal cell function. It is reported that, under low temperature and high pressure, the degree of unsaturation and branching of fatty acids in cell membranes is increased (DeLong & Yayanos 1985).

So far, scientists have made great progress understanding how marine bacteria regulate the degree of unsaturation in their cell membrane, but how PUFAs are generated *de novo* in marine bacteria still needs to be confirmed. However, a possible metabolic pathway for their production has been suggested (Russell & Nichols 1999). Like other organisms, bacteria generate fatty acids via a series of organized enzymes (or multi-enzyme complex), known as fatty acid synthases (FAS). In most bacteria the enzyme activities are usually arranged around a central acyl carrier protein (ACP) molecule. Covalently bound to ACP, a set of enzymes (position-specific desaturases and elongases) catalyzes the PUFA precursor, acyl-CoA, through cycles of condensation, reduction, dehydration and reduction. For example, after 7 rounds of 2C-unit addition from malonyl-ACP, acetyl-ACP will be converted to palmitic acid (C16:0). Depending on the need of the
bacteria, there is a functional distinction given by the bacterial FAS based on the different activated dehydrases. Bacteria can synthesize saturated fatty acids in a similar way to the systems employed by animals, plants and eukaryotic microorganisms, or produce a mixture of saturated and monounsaturated products. As shown in Fig. 1.3.2, because of the different thermal stabilities between the saturated and unsaturated intermediates, bacteria can balance these two pathways. In this case, more unsaturated products will be generated to help bacteria survive at a lower temperature. None of these reactions require the direct involvement of oxygen. Thus, it is also known as the “anaerobic pathway”.

Since the anaerobic pathway could only introduce one double bond per fatty acid (implying that they have only a single type of desaturase/dehydrase), a mechanism requiring a series of desaturases, elongases as well as molecular oxygen may be needed (Nichols & Russell 1996; Yazawa 1996). As the double bonds introduced during desaturation are known to be position-specific relative to the carboxyl terminus because of the specificity of desaturases (Quint & Fulco 1973), to complete the desaturation, unless desaturases with other positional specificities are involved, there has to be an elongation system (or a set of specific elongases) as well. However, by using radiolabelled malonyl-CoA, separate desaturation and elongation reactions for the EPA biosynthesis pathway were not identified. The result indicates that the complete pathway from FAS primer to final PUFAs might be under the regulation of an integrated complex.
Fig. 1.3.2 “Anaerobic” fatty acids biosynthesis in Bacteria Steps of the pathway

a) From 1 to 3, FAS produce only saturated fatty acids; b) From 1 to 4, FAS produce a mixture of saturated and unsaturated fatty acids.

1.3.2 Polyketide synthase (PKS) pathway in marine bacteria

The discovery of how marine bacteria produce PUFAs occurred in 1996 by the Japanese scientist Yazawa, who first successfully cloned a gene cluster that was involved in the biosynthesis of EPA from *Shewanella pneumatophori* SCRC-2738 and finally obtained recombinant bacterial EPA in *E. coli* (Yazawa 1996). This 38-kb EPA synthesis gene cluster was first cloned as a recombinant cosmid. Among all the 18 open reading frames (ORFs) carried by the inserted fragment, only five were necessary for the biosynthesis of EPA, namely *pfaA*, *pfaB*, *pfaC*, *pfaD*, and *pfaE*. Following the development of bioinformatics and genetics, scientists were able to identify 11 functional domains in these 5 ORFs. Gene sequencing and homology analysis implied that only 3 domains show similarity to FAS protein domains, while the others (8 domains) are homologous to
polyketide synthase (PKS) proteins (Metz 2001).

Polyketide synthases are normally encoded by certain kinds of bacteria, fungi, and plants that synthesize polyketides. Similar to the fatty acid synthesis catalyzed by FAS, PKS synthesizes its final product from acyl-CoA by repeated condensation and decarboxylation. PKSs are normally multi-domain enzymes/complexes. Take type I PKS for example, it contains acyltransferase (AT), which binds to starter groups, and ketoacyl-ACP synthase (KS) to elongate the polyketide chain. Optionally, depending on the purpose, dehydrase (DH) can split H$_2$O to form $\alpha$, $\beta$-double-bonds, and enoyl reductase (ER) can reduce $\alpha$, $\beta$-alkene to methylene during synthesis (Jenke-Kodama et al. 2005).

Based on genetic analysis, the gene cluster found in *Shewanella* encodes malonyl-CoA:ACP acyl transferase, 3-ketoacyl synthase, 3-ketoacyl reductase, acyltransferase, pantetheine transferase, chain elongation factor and a cluster of ACP-like domains. Along with the other 3 domains that show homology to bacterial FAS (one is similar to Triclosan-resistant enoyl reductase and the other two dehydrases are homologs of *E. coli* FAS proteins encoded by gene *fabA*), researchers sketched a hypothetical polyketide synthase-based EPA synthesis pathway in *Shewanella*. Unlike any previously classified PKS proteins, the *Shewanella* PKS pathway suggested that, besides repeated processes to form the fatty acid chain, it also has the ability to isomerize the fatty acid chain in certain cycles, and with the FabA dehydrase-isomerase and Triclosan-resistant enoyl reductase involved, this gene cluster shows biochemical similarity to both PKS and FAS (as shown in Fig.1.3.2.1).
1.3.3 The PKS pathway found in *Shewanella baltica* MAC1

As mentioned above, fatty acids in bacteria can be synthesized via the FAS or PKS pathway. Because several different reactions/enzymes are involved in polyunsaturated fatty acid synthesis, it indicates that two separate sets of enzymes encoded by different genes/gene clusters are responsible for catalyzing the reactions in FAS or PKS.

**Fig. 1.3.2.1** Putative EPA biosynthesis pathway by *Shewanella* polyketide synthase. DH, dehydrase; ER enoyl reductase; I, trans, cis isomerase; KR, 3-ketoacyl reductase; KS, 3-ketoacylsynthase (Wallis et al. 2002).
Compared to the classic FAS pathway, the PKS pathway requires no oxygen and less energy. Thus, it allows marine bacteria to produce more long chain polyunsaturated fatty acids to survive in extreme environments (low temperature and high pressure).

More recently, great enhancement of omega-3 fatty acid production has been achieved by randomly mutated *Shewanella baltica* MAC1 mediated by a Tn5 transposon (Amiri-Jami et al. 2006). The Tn5-derived transposons are designed to generate random mutations (de Lorenzo et al. 1990). It contains an antibiotic resistance gene for selection, a unique NotI cloning site and a 19 base-pair terminal repeat sequence of Tn5 as a flanking region. By screening over 500 mutants, three were found to produce EPA at levels up to 5-times more than the wild type *Shewanella baltica* MAC1. Moreover, results showed that the EPA biosynthesis enzymes might be affected by the insertion of mini-Tn5. Unlike the wild type *Shewanella baltica* MAC1, the mini-Tn5 inserted *Shewanella* strains were able to produce EPA even at 30°C (Amiri-Jami et al. 2006).

With the discovery of these strains, Amiri-Jami et al. (2010) started to focus on cloning this EPA/DHA producing gene cluster. They prepared a fosmid library from the genomic DNA of *Shwanella baltica* MAC1 and used a partial fragment of its *pfaA* and *pfaD* genes as probes to screen EPA- and DHA-related genes by colony hybridization of a clone library maintained in *E. coli*, in an attempt to isolate this gene/gene cluster. As a result, a 35-kb gene cluster was identified and cloned into selected *E. coli* strains. Gas chromatography (GC) and gas chromatographic-mass spectrometry (GC-MS) analysis also confirmed EPA/DHA production by the recombinant *E. coli* from parent strains EPI300T1, DH5α, DH10B and JM110.
With the whole sequence of the 35-kb DNA fragment confirmed, a total of 16 potential ORFs were suggested. Analysis showed that 5 ORFs with a total size of 20 kb were homologous with *pfaE* (*ppTase*), *pfaA*, *pfaB*, *pfaC* and *pfaD* genes in the GenBank nucleotide database. However, 7 out of 16 ORFs did not show any significant similarity to GenBank sequences, and thus were considered as hypothetical proteins. In addition, eight enzyme domains in *pfaA* were also identified by domain analysis. The result suggested that functional domains in *pfaA* might contribute to the bacterial polyketide synthase-type multienzyme complexes in producing bacterial EPA/DHA (shown in Fig. 1.3.3).
Fig. 1.3.3 *Shewanella baltica* MAC1 EPA/DHA gene cluster and its flanking region. (From Amiri-Jami & Griffiths 2010) The DNA sequencing and the BLAST results (comparing the isolated 35-kb EPA/DHA gene cluster to the identified *pfa* genes obtained from GenBank) indicated that, out of the 16 potential predicted ORFs, ORF4 (1004 bp), ORF5 (8114 bp), ORF6 (2250 bp), ORF7 (3367 bp) and ORF8 (1705 bp) were similar to the *pfaE* (ppTase), *pfaA*, *pfaB*, *pfaC* and *pfaD* genes, respectively. The eight prospective domains shown in *pfa* were: β-ketoacyl-acyl carrier protein synthase (KS), β-ketoacyl-acyl carrier protein (ACP), β-hydroxyacyl-ACP dehydratase (HD), acyl transferase (AT), β-ketoacyl reductase (KR), short-chain dehydrogenase reductase (DR), chain length factor (CLF) and phosphopantetheinyl transferase (ppTase).

However, surprisingly, a high degree of homology was not found between the *pfaE* of *S. baltica* MAC1 and *S. pneumatophori* SCRC-2738, and no homology was found between the *pfaB* isolated from *S. baltica* MAC1 with the *pfaB* isolated from other marine bacteria, which was reported as a key enzyme determining whether the final product resulting from PKS pathways is EPA or DHA (Orikasa et al. 2009). The domain analysis suggested that proteins/enzymes encoded by the isolated *pfa* gene have similar/same functions to those present in the bacterial PKS pathway.
Therefore, the study suggested a new PKS pathway was responsible for PUFA biosynthesis in *S. baltica* MAC1. Excitingly, the production of EPA/DHA in the transgenic *E. coli* strains was significantly higher than that in wild type MAC1. Furthermore, by deleting the predicted unnecessary EPA/DHA genes, the new 20 kb EPA/DHA gene cluster showed even greater potential in bio-synthesizing EPA and DHA (details are in next section).

### 1.3.4 Heterologous expression of the EPA/DHA biosynthesis gene cluster found in *Shewanella baltica* MAC1

To isolate the novel, efficient bacterial EPA/DHA biosynthesizing gene cluster, a series of heterologous expression steps were followed (Amiri-Jami et al, 2010; Amiri-Jami et al, 2014). When carried by vector plasmid pCC1 FOS, the 35 kb EPA/DHA gene cluster can be stably maintained in *E. coli* strains. Gas chromatography showed that the recombinant production of EPA in *E. coli* EPI300T1 was significantly higher than *S. baltica* MAC1 at 10°C, 15°C, 20°C and 25°C. Also, the recombinant *E. coli* EPI300T1 could produce DHA even at 20°C while *S. baltica* MAC1 was only able to produce DHA at temperatures below 10°C. Because this unique EPA/DHA producing gene cluster can be easily isolated from the in-stock fosmid vector, it can be manipulated using molecular methodologies (Amiri-Jami & Griffiths 2010).

To optimize the production of the recombinant omega-3 fatty acid and to introduce this effective EPA/DHA producing gene cluster into food-grade microorganisms, Amiri-Jami et al, (2014) deleted the unnecessary flanking ORFs of the EPA/DHA gene cluster, and thus made it even more effective; 3.5- to 6.1-fold more recombinant EPA was detected from the recombinant *E. coli* strains transformed with the 20-kb gene cluster than those
carrying the 35-kb DNA fragment. Following this, the 20-kb gene cluster was successfully ligated to a low copy number vector, pIL252m (4.7 kb, Ery), and expressed in *Lactococcus lactis* subsp. *cremoris* MG1363. The confirmation of the production of recombinant EPA and DHA from *L. lactis* demonstrated the first successful cloning of an EPA/DHA gene cluster from a marine bacterium to lactic acid bacteria (Amiri-Jami et al. 2014).

Heterologous expression of target genes to produce bioactive product has always been a research hot spot. In the pharmaceutical industry, genetically modified bacteria/yeast strains are widely used in producing compounds like insulin, erythropoietin and hepatitis B vaccine to help millions of people defeat disease. In the food area, although faced with opposition, the application and promotion of GMO is progressing (details are in section 1.5). As important nutrients, so far, marine fish are still our primary source of EPA and DHA. However, the seasonal and climatic instability of the fish supply, globally decreasing fish stock, more and more serious ocean pollution, accompanied by the high cost of seafood, and the unpleasant flavor of fish oils, largely limits the production and promotion of EPA and DHA. Similarly, before bioengineered *E. coli* were applied in producing recombinant insulin, it was either extracted from the pancreas of cattle and pigs and purified, or chemically synthesized, which takes a lot of labor and might cause adverse reactions in some cases (Ward & Lawrence 2011). There is no doubt that biotechnologies significantly boosted the production and quality of insulin, minimized financial cost, and, most importantly, helped millions of diabetic patients live longer and healthier lives. Likewise, since the gene cluster responsible for biosynthesizing EPA and DHA were isolated and successfully expressed in *L. lactis* (Amiji-Jami et al. 2014), in the
near future, people may benefit from biotechnology again if we can fully develop the potential of the 20 kb EPA/DHA biosynthesis gene cluster to find a safe, easy and low-cost way to produce edible and sensorially acceptable EPA and DHA. Historically, LAB were used by human beings in food fermentation and preservation. So far, a genetically engineered *Lactococcus lactis* strain secreting the anti-inflammatory cytokine IL-10 has been tested in clinical trials and showed a clinical benefit to patients suffering from Crohn’s disease, a chronic inflammatory bowel disease (Piñero-Lambea et al. 2015). Thus, with scientific support, once applied in food industrial fermentation, genetically engineered LAB will greatly improve our quality of life.

1.4 *Streptococcus thermophilus*

1.4.1 A brief introduction of *Streptococcus thermophilus*

*Streptococcus thermophilus* is a gram-positive, facultatively anaerobic bacterium that is frequently isolated from dairy environments (Thunell & Sandine 1985). Taxonomically, based on DNA-DNA re-association, *Streptococcus thermophilus* was proved to serve as a full species in the salivarius group of viridans streptococci (Facklam 2002; Hols et al. 2006). Moreover, phylogenetically speaking, comparative whole genome analysis and concatenated alignments of ribosomal proteins revealed that *Streptococcus thermophilus* is also related to *Lactococcus lactis* (Makarova et al. 2006). Commercially, it plays an important role in yoghurt and cheese manufacture as a starter culture. The annual dairy product market value produced using *S. thermophilus* can reach as high as $40 billion US (Bolotin et al. 2004). Thus, it is considered as the second most important dairy starter after *Lactococcus lactis*. 
In the dairy industry, *S. thermophilus* is a thermophilic lactic acid bacterium (LAB) widely used as a starter. Because *S. thermophilus* is able to grow or survive at the high temperatures (45 °C) required in several production processes, it is traditionally applied to the production of yogurt and cheese, such as the hard, cooked cheeses, Mozzarella and Cheddar. For cheese-making, *S. thermophilus* is used alone or in combination with several lactobacilli starters, but for yogurt it is always used with *Lactobacillus delbrueckii* subsp. *bulgaricus* (Auclair & Accolas 1983).

Moreover, its ability to rapidly decrease the pH of milk during fermentation by promptly converting lactose into lactic acid also makes *S. thermophilus* important for industrial applications. Besides, *S. thermophilus* can produce beneficial metabolites including i) exopolysaccharides (EPS) (Broadbent et al. 2003); and ii) several bacteriocins against *Pediococcus acidilactici*, *Clostridium tyrobutyricum*, *Clostridium sporogenes*, *Clostridium botulinum*, *Bacillus cereus* and *Listeria monocytogenes*. (Aktypis & Kalantzopoulos 1998; Gilbreth & Somkuti 2005; Mathot et al. 2003).

Taxonomically, *S. thermophilus* was once considered as a subspecies of *Streptococcus salivarius*, however, multilocus sequence typing (MLST) showed that *Streptococcus thermophilus* does not share alleles with *S. salivarius* nor *S. vestibularis* in the Viridans streptococci salivarius group. Therefore, it can be classified as a distinct species (Hols et al. 2006).

It is widely accepted that *Streptococcus thermophilus* is a “Generally Recognized as Safe” (GRAS) organism. However, some species in the genus *Streptococcus*, such as *Streptococcus pyogenes* and *Streptococcus pneumonia*, are pathogenic. Thus, it is
possible that \textit{S. thermophilus} has the potential to be pathogenic. Recently, the whole genome of \textit{S. thermophilus} was sequenced. Strikingly, comparative genomic analysis showed that the virulence-related genes characterized/reported previously in pathogenic streptococci are either absent or are silenced/inactivated in \textit{Streptococcus thermophilus}. (Bolotin et al. 2004). This strongly supports the safety status of \textit{S. thermophilus}.

1.4.2 \textit{Streptococcus thermophilus} as probiotic

There are over 100 trillion microorganisms carried in the human digestive tract, which is over 10 times greater than the total number of human cells in the body, and 99% of them are anaerobic (Guarner & Malagelada 2003). These microorganisms are essential to build up a healthy and strong intestinal tract.

It has been reported that a probiotic mixture containing \textit{Streptococcus thermophilus} has a health-promoting influence on human beings, especially on maintaining the wellness of the gastrointestinal tract. Studies show that the probiotic mixture VSL#3, which contains \textit{S. thermophilus}, can effectively relieve the symptoms of ulcerative colitis (Bibiloni et al. 2005). Oral supplementation of probiotics (\textit{Bifidobacteria infantis, Bifidobacteria bifidus} and \textit{Streptococcus thermophilus}) can significantly reduce the incidence and severity of necrotizing enterocolitis in premature infants (Bin-Nun et al. 2005), and together with \textit{Bifidobacteria bifidus}, supplementation of \textit{Streptococcus thermophilus} can help prevent antibiotic-associated diarrhea and acute diarrhea in infants (Corrêa et al.; Ezez. 2011).

It is also reported that certain \textit{Streptococcus thermophilus} strains can secrete biosurfactants to interfere with the adhesion of some pathogenic species. For example, \textit{S. thermophilus} B can release biosurfactants to prevent the adhesion of \textit{Candida} spp.
(Busscher 1997). *S. thermophilus* NCC1561 can grow in dental plaque biofilm with supra- gingival plaque forming bacteria, and thus interfere with the adhesion of these oral species (Comelli et al. 2002).

With the increasing awareness of the concept of a healthy lifestyle, the global market for functional foods and nutrient supplements has expanded rapidly in the last decade. Probiotic supplements generated revenues of $15.9 billion US in 2008 and were expected to be worth $32.6 billion US by 2014 with a compound annual growth rate of 12.6 percent from 2009 (Ganguly, 2011). However, unexpectedly, the actual revenue achieved reached $58.7 billion US by the year 2013, and is now expected to reach $96 billion US by 2020. So far, several strains of *Streptococcus thermophilus* have been commercialized and used in daily life. For example, helping improve lactose digestion in the intestinal tract, *Streptococcus thermophilus* ST21 is present in FloraFIT® (by DuPont™ Danisco®) as one of the 15 probiotic strains for customers to customize their own probiotic supplement.

**1.5 Prospects of genetically modified (GM) food**

Nowadays, the use of recombinant DNA biotechnological procedures allows us to genetically alter organisms in the way we need, and thus result in the appearance of genetically modified (GM) food and genetically modified organisms (GMO). By moving genes from one organism to another or by deleting/altering the original genes in the organisms result in the expression of novel attributes that did not exist before. In fact, the commercialization of GM food and GMOs has been at the center of vehement controversy since the day they were developed. So far, two major points of view are held by society; on one hand, DNA engineering is considered a potential tool to significantly
enhance food/therapeutic productivity and quality. For example, *Bacillus thuringiensis* maize (also known as BT corn, a bioengineered corn transformed with *Bacillus thuringiensis* genes that can produce Bt delta endotoxin, which acts as a pesticide) is resistant to European corn borer (Romeis et al. 2006); rennet produced by genetically modified microorganism are approved for use in cheese making (Snow et al. 2005); and in therapeutics, genetically modified *E. coli* is widely applied in the production of human insulin (Williams et al. 1982).

On the other hand, opponents argue that the effects of the introduction of these engineered genes carried by GMOs and GM food to the environment and human beings are still unknown, or even might be harmful. Firstly, the recombinants have an obvious survival advantage, thus, they might dominate the native population, and thus become the main community in the environmental niche, resulting in great loss of biodiversity. Moreover, it is reported that cloned genetic material can be transferred from GM plants into related non-GM species in the wild, and there is also concern about gene transfer from GM food or GMOs to human cells or their gut microflora (Buiatti et al. 2012).

Legal and societal approval are imperative for the widespread use of GMOs and GM food. Government attitudes towards genetic modification differ widely, especially in North America and Europe. Take labeling of GM foods for example, in Europe it is required that all food products that make direct use of GMOs at any point in their production be labelled as such, regardless of whether the GM content can or cannot be detected in the final product. Also, any food or feed containing more that 0.9 % of GM content must be labelled (Regulations 1829/2003 and 1830/2003) (Davison 2010). However, North America has a relatively loose legal environment towards the GMO and GM foods
labelling issue. In the USA and Canada, it is not mandatory to label GM food, unless the transgenic food is materially different from its conventional counterpart (FDA, 1992; Canadian General Standards Board, 2004).

In addition, consumer opinions are also important to the future of GMOs and GM foods. Consumer concerns over these novel foods are basically due to information asymmetry, incomplete information and, to a certain extent, the strong opposition by activist groups (which may or may not be driven by politics) as well as the persistent advocacy of negative opinions by the mass media. As a matter of fact, to remove opponents’ suspicions, studies around GMOs side effects on the environment and human health have never stopped. Although it is true that gene flow does happen between GM crops and their wild relatives; a 10-year study in the UK reported that GM corn, oilseed rape, potato and sugar beet are no more invasive or persistent than their conventional counterparts (Crawley et al. 2001). Besides, molecular biological methods were also developed to minimize the occurrence of gene flow by developing systems for selectable marker excising (Hare & Chua 2002). Since the first approved GM crop was largely cultivated in 1996, not a single case of adverse reaction (poisoning or allergic) caused by the consumption of GM food products was reported. Further more, several studies indicating adverse effects of GM food in animal tests were found to be biased. For example, recently it was reported that microRNA from plants (rice) can accumulate in mammalian blood and tissues and might be involved in mammalian gene regulation (Zhang et al. 2012); with this knowledge, GMO-opponents grossly exaggerated the influence of this phenomenon. However, GMO-opponents neglected a few facts. Above all, in this research, the amount of raw rice ingested to the rats daily is equivalent to 33 kg of rice
per day for human being, which will never happen in reality. Not to mention the fact that exogenous microRNA is extremely difficult for the human body to absorb orally (the pharmaceutical industry is still trying to stabilize microRNAs as an oral medication to suppress cancer cells) (Dickinson et al. 2013). In addition, microRNA is a form of genetic material that widely exists among all plants and animals, therefore, even without the existence of GM food, we still consume large amounts of plant/animal microRNAs in our everyday diet.

Therefore, under the control of strict laws and regulations based on scientific facts, with a trustworthy and comprehensive third party risk assessment, accompanied by proper and scientifically sound advertising, recombinant DNA technology may be a reliable “game changer” in the food industry that can help human beings survive by solving the problem of famine, and may improve our quality of life by bringing us more nutrient-rich food with beneficial health properties.

1.6 Research objectives

Considering the stability and high-level expression of the EPA/DHA-producing gene in *E. coli* transformants as well as the previous successful construction of EPA/DHA-positive *L. lactis* subsp. *cremoris* MG1363 clones, we hypothesized that 1) the 20-kb EPA/DHA gene cluster will be stably maintained and expressed in *E. coli*, even without the supplementation of antibiotic as selection pressure; and 2) *S. thermophilus* ST21 can be transformed with the 20-kb EPA/DHA-producing gene cluster from *Shewanella baltica* MAC1 and thus produce recombinant EPA and/or DHA.

Thus, in this project, our research objectives are
1) To test the stability of the 20-kb EPA/DHA gene cluster in consecutive mass cultivation (with periodic sub-culture) of recombinant *E. coli* EPI300T1 strain no.6 without antibiotic supplementation and to analyze the recombinant production of EPA/DHA.

2) To genetically modify *S. thermophilus* ST21 to generate a novel *S. thermophilus* strain that can produce recombinant EPA and DHA and analyze the stability of the inserted foreign DNA fragment and the recombinant production of EPA and DHA.
2. Stability of EPA/DHA production by recombinant *E. coli* EPI300T1 no. 6 harboring plasmid pfBS-PS in consecutive cultivation

2.1 Introduction

In the past two decades, scientists have tried to characterize and isolate effective EPA/DHA genes found in marine bacteria in order to transfer them to other species (Metz 2001; Nishida et al. 2006; Okuyama et al. 2007). Previously, our research group identified a 35 kb EPA/DHA-producing gene cluster in *Shewanella baltica* MAC1, and successfully transformed five *E. coli* strains with this marine-derived EPA/DHA gene cluster (Amiri-Jami & Griffiths 2010). More recently, Amiri-Jami et al. (2014) genetically modified the 35 kb EPA/DHA gene cluster by deleting its unidentified flanking DNA fragment. A series of recombinant *E. coli* strains harboring the modified 20 kb EPA/DHA gene cluster cloned from *Shewanella baltica* MAC1 was constructed (Amiri-Jami et al. 2014).

Among all the *E. coli* strains mentioned above, *E. coli* EPI300T1 harboring plasmid pfBS-PS effectively produced EPA and/or DHA under antibiotic pressure (Amiri-Jami et al. 2014). Plasmid pfBS-PS was constructed based on the commercialized, low copy number, dephosphorylated vector pCC1FOS (Epicentre). In pfBS-PS, chloramphenicol resistance was used as the selection marker (Fig. 2.1).
The repaired 20 kb EPA/DHA gene cluster was ligated to the Eco72 I (blunt) site of the dephosphorylated, CopyControl pCC1FOS vector. The NotI restriction site is shown (From Amiri-Jami et al. 2014).

It is widely accepted that the genetic background of the host cell, the copy number of the vector plasmid, the size of the insert DNA fragment as well as culture conditions, especially the environmental selection pressure applied, can significantly affect the stability of the recombinant plasmid (Smith & Bidochka 1998).

Considering that the size of plasmid pfBS-PS (8 kb pCC1 FOS + 20 kb EPA/DHA gene cluster) is relatively large, and the fact that recombinant production of EPA was only tested in the presence of the antibiotic to provide selective pressure, it is important to determine whether pfBS-PS can stably exist in the recombinant *E. coli* EPI300T1 and
produce recombinant EPA/DHA without the presence of antibiotic in the media. As a result, in this chapter, the stability of the 20-kb EPA/DHA gene cluster and the recombinant production of EPA/DHA in *E. coli* EPI300T1 were examined by subjecting the recombinant strain to successive sub-culturing every 4 days for 20 days with or without chloramphenicol selection pressure.

### 2.2 Materials and methods

#### 2.2.1 Bacterial strains and culture conditions

Stock cultures of *E. coli* EPI300T1 no.6 were maintained frozen at -80°C in 15% glycerol. Bacterial cultures for use in experiments were prepared by inoculation of frozen stock cultures onto Luria Bertani agar (LBA) (Difco, Detroit, USA) plates supplemented with 12.5 µg ml⁻¹ chloramphenicol (Sigma-Aldrich, St. Louis, MO, USA). Plates with *E. coli* EPI300T1 no.6 were incubated at 37°C overnight. To extract the plasmid DNA of pfBS-PS (8 kb pCC1FOS+ 20 kb EPA/DHA gene cluster, Cmᵀ), an isolated colony from an overnight culture was inoculated in 50 ml tubes containing 10 ml Luria Bertani (LB) (Difco, Detroit, USA) broth supplemented with 12.5 µg ml⁻¹ chloramphenicol at 37 °C overnight with shaking at 200 rev min⁻¹.

Two hundred microliters of the confirmed *E. coli* EPI300T1 no.6 inoculum were then added aseptically to 250 ml LB broth supplemented with 12.5 µg ml⁻¹ chloramphenicol and labeled as *E. coli* EPI+. In the meanwhile, the same amount of *E. coli* EPI300T1 no.6 inoculum was also added aseptically to each of 3 flasks containing 250 ml LB broth without the supplementation of chloramphenicol, and labeled as *E. coli* EPI1, EPI2, and EPI3, respectively. All growth experiments were done independently at 15 °C and
cultures sampled to analyze for EPA/DHA production and to perform plasmid extraction every 4 days for up to 20 days.

**Table. 2.2.1 Bacterial strains used in this study**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> EPI300T1 no. 6</td>
<td>EPI 300T1 harboring pfBS-PS</td>
<td>CRIFS collection (Amiri-Jamie al, 2010)</td>
</tr>
<tr>
<td><em>E. coli</em> EPI1+</td>
<td><em>E. coli</em> EPI300T1 no. 6 (12.5 µg ml(^{-1}) chloramphenicol)</td>
<td>This study</td>
</tr>
<tr>
<td><em>E. coli</em> EPI1</td>
<td><em>E. coli</em> EPI300T1 no. 6 cultured in LB broth for 20 days (sub-cultured every 4 days) without supplementation of chloramphenicol in series 1 flasks</td>
<td>This study</td>
</tr>
<tr>
<td><em>E. coli</em> EPI2</td>
<td><em>E. coli</em> EPI300T1 no. 6 cultured in LB broth for 20 days (sub-cultured every 4 days) without supplementation of chloramphenicol in series 2 flasks</td>
<td>This study</td>
</tr>
<tr>
<td><em>E. coli</em> EPI3</td>
<td><em>E. coli</em> EPI300T1 no. 6 cultured in LB broth for 20 days (sub-cultured every 4 days) without supplementation of chloramphenicol in series 3 flasks</td>
<td>This study</td>
</tr>
</tbody>
</table>

### 2.2.2 Test of the stability of plasmid pfBS-PS in *E. coli* EPI300T1 no.6

### 2.2.2.1 Extraction of plasmid

A single colony selected from an LB agar plate supplemented with 12.5 µg ml\(^{-1}\) chloramphenicol was transferred to 10ml LB broth with chloramphenicol, and incubated
at 37°C overnight for plasmid extraction. The extraction of plasmids was conducted with QIAprep Spin Miniprep Kit (Qiagen, Toronto, Canada). One milliliter of each sample (overnight culture) was added to an autoclaved 1.5 ml microcentrifuge tube. The tubes were centrifuged at 15,000 × g for 3 min in a microfuge (Beckman Coulter, Mississauga, ON) to pellet the cells. To obtain a higher plasmid concentration, this step was repeated to collect cells from a total volume of 5 ml of the overnight culture.

To remove RNA, 300 µl buffer P1 (mixed with RNAse) were added and mixed thoroughly with the pelleted cells. After cells were completely suspended, 300 µl of buffer P2 were mixed gently to lyse the cells at room temperature for 5 min. Four hundred microliters of buffer N3 were added, mixed thoroughly and immediately to stop the lysis reaction and denature proteins in the cells. Samples were then centrifuged (12 min, 15,000 × g) and the supernatant was transferred carefully to the QIAprep spin column by pipetting.

In the QIAprep spin columns, the transferred solution was again centrifuged (1 min, 15,000 × g) and the flow-through was discarded. The QIAprep spin columns with the retained plasmid DNA were washed with 500 µl of buffer PB (binding buffer) and 750 µl Buffer PE (wash buffer), successively. After completely removing the binding buffer and wash buffer by centrifugation (2 min, 15,000 × g), the QIAprep spin columns were transferred to the clean collection tubes. At the end, the QIAprep spin column with retained plasmid DNA was washed with pre-heated elution buffer (EB) (10 mM Tris·Cl, pH 8.5). To better dissolve plasmid DNA, the spin columns with buffer EB were left to stand at 55°C for 2 min. Finally, plasmids contained in buffer EB were collected by centrifugation (2 min, 18,000 × g).
The plasmid DNA concentration of each sample was determined using the NanoDrop 1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) by measuring absorbance at 260_nm. In addition, the ratio of absorbance at 260_nm:absorbance at 280_nm was also calculated by the NanoDrop 1000 to determine the DNA purity. Isolated plasmid DNA was stored at 4°C.

2.2.2.2 Plasmid digestion and gel electrophoresis

Extracted Plasmid DNA was digested by NotI-HF restriction enzyme (New England Biolabs, Whitby, Canada) at 37 °C overnight (reaction setup shown in Table. 2.2.2.2). To test the stability and existence of recombinant plasmid pfBS-PS, plasmid extractions and NotI-HF digestions were applied to all independent cultures after 20 days. NotI-HF digested plasmids were separated in a 1 % agarose gel, with a Lambda-HindIII DNA ladder used as a DNA marker (Fisher Scientific, Ottawa, ON, Canada).

Low melting point agarose gel (1%) was prepared by melting 0.6 g agarose powder (Invitrogen, Burlington, ON, Canada) in 60 ml of 1× Tris-Acetate-EDTA (TAE) buffer. The mixture was boiled in a microwave until a clear solution was obtained. The solution was cooled to 40-45°C, 0.5 µl of 10 mg ml⁻¹ ethidium bromide (Sigma-Aldrich, St. Louis, MO, USA) were added and mixed thoroughly. Then the solution was poured into a casting stand and solidified at room temperature (23°C). Extracted plasmid/digested samples and markers were placed into the wells. The gel was placed into 1× TAE buffer in the gel holding template in the electrophoresis chamber. A constant voltage of 80 V was applied to get a good separation of DNA. The gel electrophoresis results were analyzed using the Bio-Rad Gel Doc system (BioRad, Mississauga, ON, Canada).
**Table 2.2.2.2. Restriction enzyme digestion setup**

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid DNA</td>
<td>200 ng</td>
</tr>
<tr>
<td>NotI-HF</td>
<td>1.2 µl</td>
</tr>
<tr>
<td>CutSmart® Buffer</td>
<td>2.1 µl</td>
</tr>
<tr>
<td>H$_2$O</td>
<td>Based on the plasmid DNA concentration</td>
</tr>
<tr>
<td>Total volume</td>
<td>20 µl</td>
</tr>
</tbody>
</table>

**2.2.3 Total lipid extraction of recombinant E. coli EPI300T1**

The harvested bacterial cells were first centrifuged and washed twice with de-ionized water and the cells were obtained as freeze-dried cultures in vacuum-sealed ampoules at -30°C. Total lipids from freeze-dried cells were extracted using the method described by Amiri-Jami & Griffiths (2010). Briefly, 200 µl of inoculum were added aseptically to each flask containing 200 ml media. All cultures were incubated at 15°C for 4 days, with shaking at 200 rev min$^{-1}$. The harvested cells were centrifuged in a Beckman J2-MC centrifuge (Beckman Instruments, Palo Alto, CA, USA) at 10,000 g for 15 min at 4°C, and then washed with distilled water. The centrifugation-washing step was repeated twice before samples were freeze-dried. All cells were freeze-dried overnight, ground to a fine powder and weighed. Four milliliters of 12% boron trifluoride in methanol were added to freeze-dried, finely ground cell powder in screw cap test tubes and heated in a boiling water bath for 15 min (Aldrich Chemical Co., Inc, Milwaukee, WI, USA) to esterify the samples. After samples were cooled down to room temperature, 4 ml of a saturated solution of NaCl were added to each tube with 1 ml of hexane (Fisher Scientific, Nepean, ON, Canada), and mixed thoroughly. Samples were centrifuged at 2,000 × g for 5 min to separate the hexane layer. The hexane layer was then transferred to a separate, clean
screw cap tube. To optimize the total lipid extraction, 1ml hexane was again mixed with the samples in the old tubes to re-extract the fatty acid methyl ester. Then, 2 ml of the hexane layer were transferred to a flask and evaporated under nitrogen. Following evaporation, 200 µl of hexane were added. All the hexane-resolved total lipid extraction samples were transferred immediately to 2 ml amber, screw thread vials with inserts and closures (Thermo Scientific, Waltham, MA, USA) and were stored at -20 °C before analysis by gas chromatography.

### 2.2.4 Gas chromatographic (GC) analysis of fatty acids

The fatty acid methyl esters were analyzed by an automated Agilent 6890 GC system (Agilent, Palo Alto, CA, USA). A fused silica SP-2560 Capillary column (100 m · 0.25 mm i.d.) and a flame ionization detector were used to separate the fatty acids. Respectively, the flow rates of the carrier gas, air and hydrogen were set at 1, 400, and 30 ml min⁻¹. The oven was programmed to increase from 110°C to 230°C, while the initial time was set at 2 min, program rate at 4 min and final time at 10 min. Compounds were identified by the comparison of relative retention time with that of pure EPA/DHA methyl ester standards (Sigma, Oakville, ON, Canada).

### 2.2.5 Statistical analysis

Two-way ANOVA was used to test significant difference between treatments (presence/absence of antibiotic and duration of incubation as factors). Groups with significantly different means were identified using Tukey's test (p < 0.05). Experiments were performed independently in triplicate.
2.3 Results and discussion

2.3.1 Stability of the recombinant plasmid pfBS-PS in recombinant *E. coli* EPI300T1

After resuscitation from the -80°C stock culture, a single colony that showed chloramphenicol resistance was randomly selected and cultured in 10 ml LB broth under antibiotic pressure at 37°C overnight prior to plasmid extraction to confirm the presence of the recombinant plasmid pfBS-PS. Because restriction enzymes can cut DNA at specific recognition nucleotides, based on the restriction site of vector pCC1FOS, *NotI*-HF can excise plasmid pfBS-PS into two DNA fragments (shown in Fig. 2.1). Thus, by knowing the size of the inserted DNA fragment and plasmid vector, restriction digestion can be used as a diagnostic method to confirm the structure of the extracted plasmid, in this case, to confirm the presence of the 20 kb EPA/DHA gene cluster insert and the 8 kb pCC1FOS vector by cutting the extracted vector at its *NotI* site.

Gel electrophoresis confirmed the presence of pfBS-PS in *E. coli* EPI300T1 no.6 (Fig. 2.3.1 A), as bands appeared with sizes of about 20 kb and 8 kb (8 kb pCC1FOS + 20 kb EPA/DHA gene cluster) (Amiri-Jami et al. 2014). Interestingly, after successive sub-culturing to a high cell density every 4 days for 20 days in LB broth supplemented with chloramphenicol, the *E. coli* EPI+ cells still maintained the 8 kb pCC1FOS + 20 kb EPA/DHA insert stably (Fig.2.3.1 B, lane e). More strikingly, even though being sub-cultured and maintained in LB broth without the selection pressure of chloramphenicol, *E. coli* EPI1, EPI2 and EPI3 also harbored the recombinant plasmid pfBS-PS (8 kb pCC1FOS + 20 kb EPA/DHA insert) as did *E. coli* EPI+ cells and *E. coli* EPI300T1 no.6 (Fig. 2.3.1 B, lane b, c, d).
Fig. 2.3.1 A *NotI*-HF digestion result of the extracted plasmid from recombinant *E. coli* EPI300 T1 no.6 (day 0) Left lane: Lambda Hind III DNA Marker; Right lane: confirmed pfBS-PS digested by *NotI*-HF.  B *NotI*-HF digestion result of the extracted plasmid (day 20) from recombinant *E. coli*.  a: λHindIII DNA Markers; b,c,d: *NotI*-HF digested plasmid extracted from day 20 *E. coli* EPI1, EPI2, and EPI3, respectively; e: *NotI*-HF digested plasmid extracted from day 20 *E. coli* EPI+.
2.3.2 Recombinant bacterial EPA and/or DHA can be stably produced in recombinant EPI300T1 *E. coli* strains without selective pressure

It has been reported that the recombinant *E. coli* EPI300T1 no.6 harboring plasmid pfBS-PS can produce EPA and DHA under antibiotic pressure (Amiri-Jami et al. 2014). Thus, in this study, we used the recombinant strain *E. coli* EPI300T1 no.6 cultured in LB media supplemented with chloramphenicol as a positive control, to discover if the recombinant *E. coli* can stably maintain such a relatively large size insert (20 kb) and express it. As a result, among all the tested strains, we were able to confirm the presence of the insert not only by gel electrophoresis but also, by using gas chromatography, we were able to detect the production of EPA by all the strains. Gas chromatograms of the putative EPA and DHA peaks were compared to the retention time of pure EPA/DHA standards. **Fig. 2.3.2.1** illustrates the EPA peak in *E. coli* EPI2 after 4 and 20 days of culture, as well as the standard EPA peak.

Similar results were obtained with the EPI+ strains, which were grown in LB broth at 15°C with the supplementation of chloramphenicol. Recombinant production of EPA could be detected in all the recombinant *E. coli* EPI300T1 mutants when grown in the presence or absence of chloramphenicol. The yield of recombinant production of EPA was also calculated (**Fig. 2.3.2.2**).
Fig. 2.3.2.1 Gas chromatographs obtained from cells of recombinants (a) Gas chromatogram of standard EPA The retention time of the standard EPA given the provided column and set method is 43.991 min; (b) Gas Chromatograph of total fatty acid methyl esters prepared from *E. coli* EPI2. Fatty acids were extracted from cells grown in LB broth without chloramphenicol for 4 days at 15°C with shaking; (c) Gas chromatograph of fatty acids extracted from *E. coli* EPI2 cells grown in LB broth without chloramphenicol for 5 successive sub-cultures over 20 days (at 15°C with shaking; (d) Total lipid extracted from *E. coli* strain with only the vector plasmid (without the 20 kb EPA/DHA insert) cultured in LB broth without chloramphenicol for 4 days at 15°C with shaking were set as a negative control. All extracted fatty acids from cells were analyzed using an Agilent 6890 GC.
The yield of the recombinant EPA (mg of EPA/ g cell dry weight) by *E. coli* EPI300T1 no.6 grown without chloramphenicol (red) and with chloramphenicol (blue).

A two-way ANOVA test was used to determine whether significant differences in EPA yield were observed between the cells grown with and without chloramphenicol. Using a confidence interval of 95%, there was no significant difference between EPA production in media with or without chloramphenicol, nor was there a significant difference between the amount of EPA produced at any time during growth for up to 20 days. Thus, the 20 kb EPA/DHA insert was maintained by *E. coli* EPI300T1 and can stably produce...
recombinant EPA even in the absence of chloramphenicol for at least 20 days with subculture every 4 days. Overall, the average yield of recombinant EPA from recombinant *E. coli* grown in LB broth without chloramphenicol was 77.74 mg g\(^{-1}\) cell dry weigh (CDW), and was not significantly different (P >0.05) from the yield provided by the same recombinant strain grown in the presence of chloramphenicol (69.99 mg g\(^{-1}\) CDW). According to the composition of the total cellular lipids, there was also no significant difference (P >0.05) in the EPA and DHA composition between the two groups. As a result, we conclude that the 20 kb EPA/DHA gene cluster can be stably maintained and expressed by the pCC1 FOS vector in its second host *E. coli* EPI300T1.

EPA and DHA concentrations in cells grown in the presence or absence of chloramphenicol are shown in **Table. 2.3.2**.

**Table. 2.3.2 The composition of EPA and DHA in total fat of the recombinant *E. coli***

<table>
<thead>
<tr>
<th>Fatty acid</th>
<th>EPI300 with chloramphenicol</th>
<th>EPI300 without chloramphenicol</th>
</tr>
</thead>
<tbody>
<tr>
<td>20:5n-3 (EPA)</td>
<td>7.604 ± 1.560</td>
<td>8.545 ± 1.227</td>
</tr>
<tr>
<td>22:6n-3 (DHA)</td>
<td>0.041 ± 0.012</td>
<td>0.021 ± 0.013</td>
</tr>
</tbody>
</table>

* Values are means of samples (% of total fatty acids) ± standard deviation (of measurement).

Although the recombinant *E. coli* strains were able to produce EPA, their ability to synthesize DHA was not as marked. Among all the samples, only a trace amount of DHA was detected by gas chromatography. However, as previously reported, recombinant DHA produced by transformed *E. coli* EPI300T1 were from traces to about 0.4% of the
total lipid extracted from freeze dried cells under different conditions (Amiri-Jami et al. 2014). In this study, the fatty acid composition observed confirmed that the yield of recombinant DHA produced by the transformants was not as high as that for EPA. Moreover, although the average percentage of DHA in total fat extracted between the two treatments varied slightly, there was no significant difference in levels between the recombinant *E. coli* grown with chloramphenicol and the strains grown without chloramphenicol (P > 0.05), which also suggested that the recombinant plasmid pfBS-PS can be stably maintained and expressed in *E. coli* EPI300T1 over 5 successive subcultures in 20 days.

2.4 General Discussion

The first successful recombinant production of EPA in *E. coli* was reported by Yazawa (1995). Later, researchers were also able to isolate *E. coli* clones capable of producing EPA and/or DHA (Orikasa et al. 2006; Amiri-Jami & Griffiths 2010). However, the stability of the recombinant production of EPA and/or DHA still needs to be tested in the absence of antibiotic before the recombinant plasmids could be used in the food industry. As mentioned above, we concluded that the 20 kb EPA/DHA gene cluster can be stably maintained and expressed by the pCC1 FOS vector in its second host *E. coli* EPI300T1.

Using genetically engineered bacteria to produce bioactive compounds/products is commonly accepted and widely applied (e.g. using genetically engineered *E. coli* to produce recombinant insulin, applying genetically modified fungi that produce recombinant rennet in cheese making). In this case, plasmids, a form of circular, double-stranded DNA, which is distinct from chromosomal DNA, is widely used as a vector in molecular cloning. Not only in laboratory research, plasmid vectors also have extensive...
industrial usage. In the food industry, the stability of recombinant plasmids during culture and fermentation is a primary concern, because plasmid loss could lead to the production of a non-active product and thus incur considerable financial loss.

As an inheritable extrachromosomal DNA structure standing independently of the bacterial chromosome, plasmids play a crucial and as well interesting role in bacterial evolution and adaptation by mediating exchanges of genetic material, letting genetic traits scatter among species, introducing mutations in certain microbial populations, and thus contributing to microbial biodiversity (Andrup et al. 2003). However, it is also the independence and the vector-host incompatibility that contribute to the instability of recombinant plasmids. Because plasmids are normally inherited independently from the bacterial chromosome, when passed on to the daughter cell only segregational-stable systems can ensure each daughter cell will receive at least one plasmid copy at cell division. It has been reported that environmental conditions, especially the lack of selective pressure, can strongly influence the segregational stability of a plasmid, and thus evoke changes to the recombinant bacterial population (Gerdes et al. 1986; Rohde 1995; Guerrier-Takada et al. 1997). In general, the characteristic given by plasmids are not essential for their hosts. Thus, only plasmids that introduce an obvious survival advantage to host cells would be maintained and passed from generation to generation stably in a competitive environmental niche. Thus, in genetic engineering, antibiotic resistance and auxotrophy are generally used as selection markers for plasmid positive clones.

It was previously reported that the recombinant *E. coli* EPI300T1 no.6 has the ability to harbor recombinant plasmid pfBS-PS (20 kb EPA/DHA gene cluster + 8 kb pCC1 FOS)
and produce a significantly high level of EPA (Amiri-Jami et al. 2014). However, the stability was only tested in medium containing chloramphenicol, and cultured for a short time (3 days). Also, according to Food and Drug Regulations – C.R.C., c.870 (Section B.01.048. (1) N and C.01.606), antibiotics are strictly regulated in the dairy industry and production agriculture. Fluid milk processors in Canada have the obligation to test raw milk before allowing tank trucks to unload milk into the plants. Any milk found to contain antibiotics is rejected by the processor. Also, the residues of antibiotics in dairy products are regulated and inspected by Health Canada and Canadian Inspection Food Agency strictly. Thus, in this study, to better simulate an industrial fermentation, we studied the stability of the pfBS-PS without any antibiotic supplementation in the medium for long term cultivation (successive sub-culture every 4 days for 20 days). The stability of the recombinant plasmid even in the absence of antibiotic survival pressure was confirmed.

The ability to stably maintain foreign DNA depends on both the characteristics of the host cells and the recombinant DNA (Smith & Bidochka 1998). Generally, as “molecular parasites”, plasmids place extra metabolic burden on the host cells. The larger the foreign DNA element is, the more complex transcription products or proteins it encodes; and thus, the heavier metabolic burden to the host cells it may cause. Thus, translating and replicating the large size foreign DNA insertion in a high copy number vector will put the host cell at a disadvantage compared to the plasmid-free bacteria and may cause plasmid segregational instability (Friehs 2003). Accordingly, a large plasmid associated with low copy number vector can help reduce the metabolic load of the host cells to improve the stability of the plasmid. In this study, the host strain, *E. coli* EPI300T1, has the
inactivated trfA gene to produce initiation factor for the high-copy replication origin oriV in the vector pCC FOS1. However, gene trfA is under the regulation of the $P_{araBAD}$ ($P_{BAD}$) promoter and AraC protein. Thus, without the induction of the expression of gene trfA, the plasmid pCC FOS1 will be kept in very low copy number. Considering the large size and the complicated cellular function of the 20 kb EPA/DHA insert, in this case, avoiding inducing the massive replication of plasmid pFBS-PS helped us obtain long term stability for recombinant production of EPA and DHA.
3 The construction of recombinant *Streptococcus thermophilus* strain harboring the 20 kb EPA/DHA gene cluster

3.1 Introduction

It has been well documented that EPA and DHA greatly contribute to human wellness by playing a role in human nutrition as “conditionally essential fatty acids”. So far, although microalgae have been used to commercially produce EPA and DHA, marine fish and fish oils are still the main source of dietary EPA and DHA, primarily because of the high cost and low productivity associated with production of these fatty acids from microalgae culture (Chi et al., 2009). However, due to the seasonal and climatic instability of the fish supply, the substantial decrease in marine fish stock, ocean pollution, the high cost of seafood, and the unpleasant flavor of fish oils, environmentally friendly, cost-effective and palatable alternative sources of these valuable marine-derived LCPUFAs are being sought.

PUFAs were once thought to be absent from bacterial membranes (Erwin, Hulanicka, and Bloch 1964). However, bacterial species of marine origin have now been shown to produce long-chain PUFAs such as EPA and DHA. It has been reported that certain cold-water marine bacteria, such as *Shewanella pneumatophori*, *Colwellia psychrerythraea* and *Shewanella baltica*, have the ability to synthesize EPA and/or DHA *de novo* (Hirota 2005; Bowman and Gosink 1998; Amiri-Jami et al. 2006). Along with developments in bioengineering and molecular biology, discovery of EPA and/or DHA producing bacteria has resulted in extensive research of their genomes. In 1996, Japanese scientists first reported the discovery of an EPA-producing gene cluster in the marine bacterium *Shewanella pneumatophori* SCRC-2738. This 38 kb gene cluster containing 9 ORFs was
later cloned into *E. coli* and resulted in EPA-producing recombinant strains (Yazawa 1996). Further sequence analysis of the 9 ORFs demonstrated that only 5 were responsible for EPA biosynthesis (Orikasa et al. 2004). Later, the same research group showed that the phosphopantetheinyl transferase (ppTase) gene cloned from the DHA-producing bacterium *Moritella marina* strain MP-1 was essential for the production of DHA.

More recently, a marine bacterium isolated from mackerel was found to have the ability to synthesize EPA and DHA, and was identified as *Shewanella baltica* MAC1 according to the comparative analysis of complete 16S rRNA gene sequences with other *Shewanella* spp. (Amiri-jami et al. 2006). Furthermore, a more productive 35 kb EPA and DHA synthesizing gene cluster in *Shewanella baltica* MAC1 was successfully isolated and expressed in *E. coli* strains, which led to a substantial enhancement of recombinant production of EPA and DHA (Amiri-Jami and Griffiths 2010). Nucleotide sequence analysis of this EPA/DHA gene cluster indicated the presence of 16 ORFs in the 35 kb sequence. According to the sequences available in GenBank, 5 (20,668 bp) out of the 16 predicted ORFs were homologous with 5 *pfa* genes (*pfaE, pfaA, pfaB, pfaC* and *pfaD*). These 5 *pfa* genes encoded 8 functional domains in enzymes involved in the PKS pathway responsible for bacterial production of EPA and DHA (Amiri-Jami and Griffiths 2010). The eight functional domains are β-ketoacyl-acyl carrier protein synthase (KS), β-ketoacyl-acyl carrier protein (ACP), β-hydroxyacyl-ACP dehydratase (HD), acyl transferase (AT), β-ketoacyl reductase (KR), short-chain dehydrogenase reductase (DR), chain length factor (CLF), and phosphopantetheinyl transferase (ppTase) (Amiri-Jami and Griffiths, 2010).
With this knowledge, to obtain even higher production of EPA and DHA, Amiri-Jami et al. (2014) downsized the EPA/DHA gene cluster from 35 kb to 20 kb by removing unnecessary flanking regions. This new 20 kb EPA/DHA gene cluster was expressed by five *E. coli* strains. As a result, the recombinant *E. coli* strains harboring the 20 kb gene cluster produced a significantly higher level of recombinant EPA than those carrying the 35 kb DNA fragment (Amiri-Jami et al, 2014). Also, as mentioned in Chapter 2 of this thesis, the recombinant *E. coli* can stably maintain the 20 kb EPA/DHA gene cluster insert without antibiotic selective pressure and produced recombinant EPA and DHA during long term cultivation.

To further discover the potential of this EPA/DHA biosynthesis gene cluster, the 20 kb DNA fragment was also successfully cloned and expressed in *Lactococcus lactis* subsp. *cremoris* MG1363, which led to the production of recombinant EPA and DHA in a lactic acid bacterium (Amiri-Jami et al., 2014).

Phylogenetically speaking, *Lactococcus lactis* and *Streptococcus thermophilus* have an intimate genetic relationship, even though they were classified as two different genera. Whole genome DNA and DNA-RNA genomic analysis of lactic acid bacteria revealed that *S. thermophilus* and *L. lactis* subsp. *cremoris* shared a very close genetic relationship. Furthermore, phylogenetic analysis of multiple protein sequences indicated that the two genera, streptococci and lactococci, shared the same branch in the *Lactobacillales* phylogenetic tree (Figure 3.1) (Makarova et al. 2006). The fact that a gene cluster originating from *S. thermophilus* could be cloned and expressed in *L. lactis* MG1363 supported the bioinformatics prediction of their similarity (Stingele et al. 1999).
*Streptococcus thermophilus* is classified as a gram-positive, facultatively anaerobic bacterium that is frequently isolated from the dairy environment. It plays an important role in the dairy industry, where it is used as a starter culture and considered as the second most important after *Lactococcus lactis*.

In recent decades, significant advances have been made in the development of techniques of DNA transformation and manipulation in microorganisms. As one of the most important LAB strains, numerous cases of successful transformation of *Streptococcus thermophilus* have been reported (Somkuti and Steinberg 1988; O'Sullivan and Fitzgerald 1999; Blomqvist, Steinmoen, and Håvarstein 2006). Also, because of the potential of genetically engineered lactic acid bacteria for industrial applications, several cloning vectors based on cryptic plasmids isolated from LAB such as *Lactobacillus*, *Streptococcus thermophilus*, and *Pediococcus* spp. have been described (Fons et al. 1997; Somkuti et al. 1998; Motlagh et al. 1994).

Based on the successful transformation of the 20 kb EPA/DHA gene cluster into *Lactococcus lactis* subsp. cremoris MG1363, the close phylogenetic relation between the two strains, and all the well-designed vector plasmids, it is reasonable to hypothesize that *Streptococcus thermophilus* strains have the ability to be transformed with the same gene cluster. In this work, we transformed the modified PIL252 carrying EPA/DHA gene cluster constructed by Amiri-Jami et al (2014) to *Streptococcus thermophilus* ST21. Then the clones positive for EPA/DHA genes were tested for EPA/DHA production.
Fig. 3.1 Phylogenetic tree of *Lactobacillales* constructed on the basis of concatenated alignments of ribosomal proteins. Species are colored according to the current taxonomy: *Lactobacillaceae*, blue; *Leuconostocaceae*, magenta; *Streptococcaceae*, red. As it is shown in the phylogenetic tree, among all the LAB strains, *Streptococcus thermophilus* has the closest relationship with *L. lactis* subsp. *cremoris* (From Makarova et al. 2006).
3.2 Material and methods

3.2.1 CRIFS stock culture confirmation

3.2.1.1 Bacterial strain and growth conditions

Stock cultures of *S. thermophilus* ST21 were maintained frozen at -80°C in 15% glycerol. Bacterial cultures for use in experiments were prepared by inoculation of frozen stock cultures onto 1% (w/v) GM17 agar (M17 agar supplemented with 1% glucose) (Oxoid Microbiology, Nepean, ON, Canada) and incubated at 37°C overnight. To extract the chromosomal DNA of *S. thermophilus* ST21, an isolated colony from an overnight culture was inoculated in a 50 ml tube containing 10 ml GM17 broth (M17 broth supplemented with 1% glucose) (Oxoid Microbiology, Nepean, ON, Canada) and incubated at 37°C overnight.

3.2.1.2 *Streptococcus thermophilus* ST21 genomic DNA extraction

The genomic DNA of an isolated colony of *S. thermophilus* ST21 was extracted using the Promega Wizard® Genomic DNA Extraction Kit (Promega, Madison, WI, USA) according to the manufacturer’s instructions. Five milliliters of an overnight culture of *S. thermophilus* ST21 were centrifuged (13,000 × g, 3 min 15°C). Cells were then re-suspended in 480µl 50mM EDTA and incubated at 37°C for 10mins before the addition of 15 mg ml⁻¹ lysozyme (Sigma-Aldrich, St. Louis, MO, USA), after which cells were again incubated at 37°C for 60 min. Cells were harvested by centrifugation (15,000 × g, 2 min). Six hundred microliters of Nuclei Lysis Solution were added and followed by 5 min incubation at 80°C, then the mixture was cooled to room temperature. RNase solution (3 µl) was added and mixed thoroughly, then cells were incubated at 37°C for 60 min. After incubation, 200 µl of Protein
Precipitation Solution were added to the solution. Samples were put in an ice-bath for 5 min and then centrifuged at 15,000 × g for 3 min. The supernatant was transferred to a clean tube containing 600 µl isopropanol at room temperature and mixed completely. The sample was again centrifuged as described above and the supernatant was decanted. Six hundred microliters of 70% ethanol were added and mixed completely to re-suspend the pellet. Samples were once again centrifuged as above. The ethanol was removed by aspiration and the DNA pellets were air-dried. The extracted DNA was re-hydrated in 50 µl Rehydration Solution for 1 h at 65°C, and then stored at 4 °C.

3.2.1.3 Polymerase Chain Reaction (PCR) procedure for house keeping gene sequence determination and PCR product sequencing

Two oligonucleotide primers (Table 3.2.1.3A) synthesized by Laboratory Services Division, University of Guelph (Guelph, Ontario, Canada) were used to amplify the housekeeping genes glcK and thrS (Bensalah et al. 2011) from the genomic DNA of Streptococcus thermophilus. Forty-five microliters of PCR SuperMix (Invitrogen Life Technologies, Burlington, Ontario) containing 22 mM Tris-HCl (pH 8.4), 55 mM KCl, 1.65 mM MgCl₂, 220 µM dNTP and 22 U recombinant Taq DNA polymerase/ml, were added to a sterile 250 µl microcentrifuge PCR tube. One microliter of 50 pmol/ µl forward primer and 1 µl of 50 pmol/ µl reverse primer were added to the 45 µl SuperMix; followed by the addition of 0.5 µl genomic DNA extraction product as a template. The final solution was gently mixed and the polymerase chain reaction was performed in a Mastercycler (Eppendorf Scientific, NY, U.S.A) under the conditions shown in Tables 3.2.1.3B1 and 3.2.1.3B2 for amplification of glcK and thrS, respectively.
The two PCR products obtained were sequenced (Lab Services Division, University of Guelph, ON, Canada) and sequences were compared to those in GenBank.

**Table 3.2.1.3A** Primers used to amplify/determine housekeeping gene *glcK* and *thrS*

<table>
<thead>
<tr>
<th>Target Housekeeping Gene Sequence</th>
<th>Primer sequence (Forward/Reverse)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>glcK</em></td>
<td>5’-TGGGCAGAAACTCAAGA-3’</td>
</tr>
<tr>
<td></td>
<td>5’-AACACCACCACCCGATAAC-3’</td>
</tr>
<tr>
<td><em>thrS</em></td>
<td>5’-ATCACTGAAGATGGGAGC-3’</td>
</tr>
<tr>
<td></td>
<td>5’-CCAAGTTTACGGGTGGA-3’</td>
</tr>
</tbody>
</table>

**Table 3.2.1.3B1** PCR conditions used to amplify *glcK* in PCR

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Duration</th>
<th>Number of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>94°C</td>
<td>4 min</td>
<td>1</td>
</tr>
<tr>
<td>94°C</td>
<td>30 sec</td>
<td>30</td>
</tr>
<tr>
<td>46°C</td>
<td>30 sec</td>
<td>30</td>
</tr>
<tr>
<td>72°C</td>
<td>2 min</td>
<td>30</td>
</tr>
<tr>
<td>72°C</td>
<td>10 min</td>
<td>1</td>
</tr>
<tr>
<td>4°C</td>
<td>Hold</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 3.2.1.3B2** PCR conditions used to amplify *thrS* in PCR

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Duration</th>
<th>Number of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>94°C</td>
<td>4 min</td>
<td>1</td>
</tr>
<tr>
<td>94°C</td>
<td>30 sec</td>
<td>30</td>
</tr>
<tr>
<td>48°C</td>
<td>30 sec</td>
<td>30</td>
</tr>
<tr>
<td>72°C</td>
<td>2 min</td>
<td>30</td>
</tr>
<tr>
<td>72°C</td>
<td>10 min</td>
<td>1</td>
</tr>
<tr>
<td>4°C</td>
<td>Hold</td>
<td>1</td>
</tr>
</tbody>
</table>
3.2.2 Test of the transformability of *Streptococcus thermophilus* ST21

3.2.2.1 Bacterial strain and growth condition

Stock cultures of *S. thermophilus* ST21 were cultured as described in Section 3.2.1.1. Stock cultures of *E. coli* DH5α harboring plasmid pGKV210 (4.4 kb with erythromycin resistance and *Eco*RI, *Sma*I, *Xma*I, *Bam*HI, *Sal*I, *Pst*I multiple cloning sites) were maintained frozen at -80°C in 15% glycerol. The *E. coli* DH5α cultures used in experiments were prepared by inoculation of frozen stock cultures onto LB agar (Difco, Detroit, USA) supplemented with 250 µg ml⁻¹ erythromycin (Sigma-Aldrich, St. Louis, MO, USA). A single colony of the pGKV 210-positive *E. coli* DH5α was picked from the selective LB agar plate and sub-cultured in 50 ml tubes containing 10 ml LB broth supplemented with 250 µg ml⁻¹ erythromycin for plasmid extraction.

Recombinant *Streptococcus thermophilus* ST21 strain harboring pGKV210 was obtained from this study. The pGKV210-positive *S. thermophilus* ST21 strains were cultured in 1% (w/v) GM17 supplemented with 2.5 µg ml⁻¹ erythromycin at 37°C overnight without shaking for plasmid extraction.

3.2.2.2 Plasmid extraction

For *E. coli* DH5α harboring pGKV 210, the plasmids were extracted as described in Section 2.2.2.1. The plasmid extraction from *Streptococcus thermophilus* was performed using a protocol similar to that for *E. coli* plasmid extraction. Because *Streptococcus thermophilus* is a gram-positive bacterium, to optimize plasmid yield, the pelleted cells were first re-suspended in buffer P1 (RNAse solution) supplemented with 15 mg ml⁻¹
lysozyme (Sigma-Aldrich, St. Louis, MO, USA), and incubated at 37 °C for 30 min. After the incubation, procedures described in Section 2.2.2.1 were followed.

3.2.2.3 Preparation of electrocompetent *Streptococcus thermophilus* ST21 cells

Electrocompetent *S. thermophilus* cells were made from highly concentrated cultures of *Streptococcus thermophilus* ST21 as follows. Before treatment, *S. thermophilus* ST21 cells were grown in 1% GM17 at 37°C overnight, the overnight culture was then diluted 100-fold in pre-heated Hogg-Jago glucose broth (HJG) (made of 3% tryptone, Oxoid Microbiology, Nepean, ON, Canada; 1% yeast extract, 0.2% beef extract, Becton, Dickinson and Company, Franklin Lakes, NJ, USA; 0.5% KH$_2$PO$_4$, 0.5% glucose Sigma-Aldrich, St. Louis, MO, USA). Cells were incubated at 37°C without shaking until the optical density at 660 nm (OD660) reached 0.3.

Ten milliliters of inoculum were added at a ratio of 1:1 to pre-warmed HJG broth containing 20% (w/v) glycine in a 50 ml tube and incubated for 1 h. Cells were then harvested by centrifugation at 4,000 × g for 10 min at 4°C in a Beckman J2-MC centrifuge (Beckman Instruments, Palo Alto, CA, USA). Cells were washed twice with 1 volume of ice-cold electroporation buffer (5 mM KHPO$_4$, 0.4 M D-sorbitol, 10% glycerol; pH 4.5). Harvested cells were re-suspended in 4 ml ice-cold electroporation buffer, divided into aliquots, and frozen in an ethanol-dry ice bath. Electrocompetent *S. thermophilus* ST21 cells were stored at -80°C prior to electroporation.

3.2.2.4 Transformation and screening for pGKV 210-positive clones

Electroporation was used to transform *S. thermophilus* using the method of Blomqvist et al. (2006) with some modifications. A Bio-Rad GenePulser XCell device (BioRad,
Mississauga, ON, Canada) was used to transform *S. thermophilus* ST21 electrocompetent cells by electroporation. Electrocompetent cells were thawed on ice, and 80 µl of the electrocompetent cell suspension were mixed with 1 - 5 µl of plasmid pGKV 210 in pre-chilled microcentrifuge tubes. The cell/plasmid DNA mixtures were incubated for 30 min on ice. After the on-ice incubation, the cell/plasmid DNA mix was transferred to pre-chilled 0.2 cm gap Gene Pulser®/Micropulser™ electroporation cuvettes (BioRad, Mississauga, ON, Canada). The GenePulser was set at 25 µF capacitance, 200 Ω resistance to deliver a single pulse of 2,500 V lasting for 5 ms. Immediately following the electroporation procedure, 1 ml of ice-cold HJGLS (HJG broth supplemented with 0.5% lactose and 0.4 M D-sorbitol) medium was added to the cell suspension. The cells were incubated at 37°C (1 - 12 h) to recover and allow the antibiotic resistant marker to express. After incubation, cell suspensions were spread on 1% GM17 agar plates supplemented with 2.5 µg ml⁻¹ erythromycin, and incubated at 37°C for 12 - 72 h. Erythromycin resistant single colonies were picked and sub-cultured for plasmid extraction and restriction enzyme digestion. Electrocompetent *S. thermophilus* ST21 transformed with no plasmid DNA were also plated on the selective medium as a negative control.

**3.2.2.5 Plasmid digestion and gel electrophoresis**

According to the structure of plasmid pGKV 210, the extracted plasmid DNA samples were digested by *BamHI* restriction enzyme (New England Biolabs, Whitby, Canada) at 37 °C overnight using the reaction mix shown in Table 3.2.2.5. To test for the presence of recombinant plasmid pGKV 210, plasmid extraction suspensions and *BamHI* digested
products were separated in 1% agarose gel using a 1 kb DNA ladder (New England Biolabs) as a marker.

The 1% agarose gels were prepared and run as described in Section 2.2.2.2.

**Table 3.2.2.5. Restriction enzyme digestion setup**

<table>
<thead>
<tr>
<th>Component</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid DNA</td>
<td>200 ng</td>
</tr>
<tr>
<td><strong>BamHI</strong></td>
<td>1.5 µl</td>
</tr>
<tr>
<td>CutSmart® Buffer</td>
<td>2.0 µl</td>
</tr>
<tr>
<td>H₂O</td>
<td>Based on the plasmid DNA volume</td>
</tr>
<tr>
<td>Total volume</td>
<td>20 µl</td>
</tr>
</tbody>
</table>

3.2.3 Transformation of *Streptococcus thermophilus* ST21 with pIL252

3.2.3.1 Bacteria strains and growth condition

Stock cultures of *L. lactis* IL1403+ harboring plasmid pIL252 (4.6 kb plasmid, with erythromycin resistance, multiple cloning sites including EcoRI, SmaI, XmaI, BamHI, SalI, PstI, low copy number) were maintained frozen at -80°C in 15% glycerol. The recombinant *L. lactis* IL1403 cultures used in experiments were cultured by inoculation of frozen stock cultures onto 0.5% (w/v) GM17 agar supplemented with 5 µg ml⁻¹ erythromycin and incubated at 30°C (Sigma-Aldrich, St. Louis, MO, USA). A single colony of the pIL252-positive *L. lactis* IL1403+ was picked from the selective GM17 agar plate and sub-cultured in 50 ml tubes containing 10 ml 0.5% (w/v) GM17 broth supplemented with 5 µg ml⁻¹ erythromycin for plasmid extraction.
Recombinant *Streptococcus thermophilus* ST21 strain harboring pIL252 was obtained from this study. The pIL252-positive *S. thermophilus* ST21 strains were cultured in GM17 broth supplemented with 2.5 µg ml\(^{-1}\) erythromycin at 37°C overnight without shaking for plasmid extraction.

### 3.2.3.2 Plasmid extraction

The extraction of plasmid pIL252 from *L. lactis* IL1403\(^+\) and the extraction of plasmid pIL252 from recombinant *S. thermophilus* ST21 were described in Section 3.2.2.2.

### 3.2.3.3 Plasmid transformation, plating, and screening the pIL252 positive clones

#### 3.2.3.3.1 Electroporation

The transformation of plasmid DNA pIL252 into *S. thermophilus* ST21 was performed as described previously (Section 3.2.2.4).

#### 3.2.3.3.2 *Streptococcus thermophilus* ST21 polyethylene glycol (PEG) mediated protoplast transformation

Also, a PEG-mediated protoplast transformation system was used to transform *S. thermophilus* ST21 (Simon et al, 1986). Cells from 10 ml of *Streptococcus thermophilus* ST21 overnight culture in 1% (w/v) GM17 were harvested by centrifugation (Beckman J2-MC centrifuge, Beckman Instruments, Palo Alto, CA, USA) at 4,000 \(\times\) g for 10 min at room temperature; followed by washing in distilled water. The washed cells were suspended in 5 ml of SMM-GM17, which was prepared by mixing equal volumes of 4 \(\times\) 1% GM17 broth and 2 \(\times\) SMM buffer (1 M sucrose, 0.04 M malic acid, 0.04 M MgCl\(_2\)). Lysozyme (Sigma-Aldrich, St. Louis, MO, USA) was added to a final concentration of 2 mg ml\(^{-1}\). The suspension was incubated for 30 min at 37°C. The protoplasts were then
pelleted at $3,000 \times g$ for 10 min and washed with 5 ml of SMM-GM17. After centrifugation ($3,000 \times g$ for 10 min), they were then re-suspended in 650 µl of SMM-GM17. Plasmid DNA (50 µl) in TE buffer (10 mM Tris HCl, 1 mM EDTA [pH 8.0]) was mixed with 50 µl of 2 × SMM buffer solution. Two milliliter of 40% (w/v) polyethylene glycol (PEG, molecular weight, 2,700 to 3,300; Sigma-Aldrich, St. Louis, MO, the USA) were added. After 2 min exposure to the PEG, protoplasts were again diluted by addition of 6.5 ml SMM-GM17 broth. Cells were pelleted by centrifugation ($3,000 \times g$, 10 min at room temperature) and re-suspended in 1.3 ml of 1% GM17 broth. To allow the erythromycin resistance gene to be expressed, protoplasts were incubated for 1 to 12 h at 30°C before being plated on 1% GM17 agar supplemented with 2.5 µg ml$^{-1}$ erythromycin.

### 3.2.3.4 Positive clone screening

The clone screening was performed as described in Section 3.2.2.4

### 3.2.3.5 Plasmid digestion and gel electrophoresis

The extracted plasmid DNA was digested by *Bam*HI restriction enzyme (New England Biolabs, Whitby, Canada) at 37°C overnight using the reaction mixture outline in Table 3.2.3.5. To test for the presence of plasmid pIL252, plasmids and *Bam*HI digested products were subjected to agarose gel electrophoresis as described in Section 2.2.2.2.

**Table 3.2.3.5 Restriction enzyme (*Bam*HI) digestion setup**

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid DNA</td>
<td>200 ng</td>
</tr>
<tr>
<td><em>Bam</em>HI</td>
<td>1.5 µl</td>
</tr>
<tr>
<td>CutSmart® Buffer</td>
<td>2.0 µl</td>
</tr>
<tr>
<td>H$_2$O</td>
<td>Based on the plasmid DNA concentration</td>
</tr>
<tr>
<td>Total volume</td>
<td>20 µl</td>
</tr>
</tbody>
</table>
3.2.4 Transformation of the plasmid pIL252m + 20 kb EPA/DHA gene cluster in

*Streptococcus thermophilus* ST21

### 3.2.4.1 Bacterial strains and growth conditions

Stock cultures of *Lactococcus lactis* subsp. *cremoris* MGED20 harboring plasmid pEDSB (pIL252m containing the 20 kb EPA/DHA gene cluster, with erythromycin resistance and *Not*I cloning site) were maintained frozen at -80°C in 15% glycerol. The recombinant *L. lactis* cultures used in experiments were cultured by inoculation of frozen stock cultures onto 0.5% (w/v) GM17 agar supplemented with 5 µg ml\(^{-1}\) erythromycin and incubated at 30°C. A single colony of the pEDSB-positive *L. lactis* was picked from the selective GM17 agar plate and sub-cultured in 50 ml tubes containing 10 ml 0.5% GM17 broth supplemented with 5 µg ml\(^{-1}\) erythromycin at 30°C without shaking for plasmid extraction.

Recombinant *Streptococcus thermophilus* STED20 strains harboring pEDSB were obtained from this study. The pEDSB positive *S. thermophilus* ST21 strains were cultured in 1% (w/v) GM17 broth supplemented with 2.5 µg ml\(^{-1}\) erythromycin at 37°C overnight without shaking.

To optimize recombinant production of EPA, *S. thermophilus* STED20 was cultivated in 1% (w/v) GM17 broth supplemented with 2.5 µg ml\(^{-1}\) erythromycin at 25°C without shaking for 3 days for cell harvest and fatty acid extraction.
3.2.4.2 Plasmid extraction
The extraction of the plasmid pEDSB from recombinant \textit{L. lactis} and the extraction of the plasmid from transformant \textit{S. thermophilus} ST21 were performed following the method described Section 3.2.2.2.

3.2.4.3 Plasmid transformation, plating, and screening the pIL252m + 20 kb EPA/DHA gene cluster positive clones
To obtain \textit{S. thermophilus} transformants with recombinant plasmid pEDSB, electroporation and \textit{S. thermophilus} polyethylene glycol (PEG) mediated protoplast transformation were performed as described in Section 3.2.3.3.

Presumptive pEDSB positive recombinant \textit{S. thermophilus} clones were screened as previously described in Section 3.2.3.3.

3.2.4.4 Colony Polymerase Chain Reaction (PCR) for \textit{pfa} gene determination
In order to detect the presence of the 20 kb EPA/DHA gene cluster, the oligonucleotide primer pair for gene \textit{pfaA}, described in Table 3.2.4.4, was used to amplify the \textit{pfaA} gene from the colonies obtained (Amiri-Jami, Lapointe, and Griffiths 2014). To lyse cells, one milliliter of overnight erythromycin-resistant transformant cultures were centrifuged (13,000 \( \times \) g, 3 min 15°C) and suspended in 150 \( \mu \)l sterile water supplemented with 15 \( \mu \)g ml\(^{-1}\) lysozyme (Sigma-Aldrich, St. Louis, MO, USA) then incubated at 37 °C for 30 min. Lysed cells were then centrifuged at 5,000 \( \times \) g for 1 min, and the supernatant collected to perform colony PCR. Forty-five microliters of PCR SuperMix (Invitrogen Life technologies, Burlington, Ontario) in 22 mM Tris-HCl (pH 8.4), 55 mM KC1, 1.65 mM MgCl\(_2\), 220 \( \mu \)M dNTP and 22 U recombinant \textit{Tag} DNA polymerase/ml, were added to a sterile 250 \( \mu \)l
microcentrifuge PCR tube. One microliter of 50 pmol/µl forward primer and 1 µl of 50 pmol/µl reverse primer were added to the 45 µl SuperMix; followed by the addition of 3 µl of lysed cell mix. The final solution was gently mixed and PCR was performed in a Mastercycler (Eppendorf Scientific, NY, U.S.A) under the conditions shown in Table 3.2.4.4B. Lactococcus lactis subsp. cremoris MG1363 harboring pEDSB was used as a positive control.

**Table 3.2.4.4A Primers used to amplify gene pfaA**

<table>
<thead>
<tr>
<th>Target Gene</th>
<th>Primer sequence (Forward/Reverse)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pfaA</td>
<td>5’-TCAGCGGTGCTGTCTGTAC-3’</td>
</tr>
<tr>
<td></td>
<td>5’-GTAGCTGCCA TGCGTATCA-3’</td>
</tr>
</tbody>
</table>

**Table 3.2.4.4B PCR conditions used to amplify pfaA**

<table>
<thead>
<tr>
<th>Temperature</th>
<th>Duration</th>
<th>Number of cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>95°C</td>
<td>10 min</td>
<td>1</td>
</tr>
<tr>
<td>94°C</td>
<td>30 sec</td>
<td>45</td>
</tr>
<tr>
<td>56°C</td>
<td>30 sec</td>
<td>45</td>
</tr>
<tr>
<td>72°C</td>
<td>2 min</td>
<td>45</td>
</tr>
<tr>
<td>72°C</td>
<td>10 min</td>
<td>1</td>
</tr>
<tr>
<td>4°C</td>
<td>Hold</td>
<td>1</td>
</tr>
</tbody>
</table>

3.2.4.5 Diagnose restriction digest of the plasmid extracted from the possible pEDSB

*Streptococcus thermophilus* positive clones

Presumptive pEDSB-positive *Streptococcus thermophilus* ST21 transformants were picked and sub-cultured at 37°C overnight without shaking in a sterile 50 ml tube
containing 10 ml 1% GM17 broth supplemented with 2.5 µg ml\(^{-1}\) erythromycin. Plasmids were extracted and analyzed by restriction enzyme digestion as described in Section 3.2.2.2

According to the structure of the recombinant plasmid pEDSB, the extracted plasmid DNA samples were digested by \textit{NotI} HF restriction enzyme (New England Biolabs, Whitby, Canada) at 37°C overnight (reaction mixture shown in Table 3.2.4.5). To test the presence of recombinant plasmid pEDSB (pIL252m + 20 kb EPA/DHA gene cluster) in \textit{S. thermophilus} clones, the \textit{NotI} HF digested products were separated in 1% agarose gel along with a Lambda-HindIII DNA ladder (Fisher Scientific, Ottawa, ON, Canada) as described in Section 2.2.2.2.

<table>
<thead>
<tr>
<th>Table 3.2.4.5. Restriction enzyme (\textit{NotI} HF) digestion setup.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmid DNA</td>
</tr>
<tr>
<td>\textit{NotI} HF</td>
</tr>
<tr>
<td>CutSmart® Buffer</td>
</tr>
<tr>
<td>\text{H}_2\text{O}</td>
</tr>
<tr>
<td>Total volume</td>
</tr>
</tbody>
</table>
3.2.5 Gas Chromatography (GC) and Gas Chromatography - Mass Spectrometry (GC-MS) analysis of the presumptive pEDSB transformed *Streptococcus thermophilus* clones

### 3.2.5.1 Total lipid extraction

Total lipid extraction from presumptive pEDSB positive-*Streptococcus thermophilus* ST21 clones, *Shewanella baltica* MAC1 and *Streptococcus thermophilus* ST21 were performed as described previously in Section 2.2.3.

### 3.2.5.2 Gas Chromatography (GC) analysis

The fatty acid methyl esters were analyzed by an automated Agilent 6890 GC system (Agilent, Palo Alto, CA, USA). A fused silica BPX 70 capillary column (60 m · 0.22 mm i.d.) and a flame ionization detector were used to separate and detect the fatty acids. Respectively, the flow rates of the carrier gas, air and hydrogen were set at 1, 400, and 30 ml min$^{-1}$. The oven was programmed to increase from 110°C to 230°C, while the initial time was set at 2 min, program rate at 4 min and final time at 10 min. Compounds were identified by the comparison of relative retention time with that of pure EPA/DHA methyl ester standards (Sigma, Oakville, ON, Canada).

### 3.2.5.3 Gas Chromatographic – Mass Spectrometry analysis

EPA/DHA production was detected by Gas Chromatographic-Mass Spectrometry (GC-MS) in the Advanced Analysis Centre (University of Guelph, Guelph, ON, Canada). EPA and DHA peaks were confirmed with an ion trap mass spectrometer. EPA and DHA were identified with target ion mass to charge ratios (m/z).
3.3 Results and discussion

3.3.1 Confirmation of the identity of *Streptococcus thermophilus* ST21

PCR result of the genomic DNA of *S. thermophilus* ST21 showed two specific bands with sizes of 500 bp and 600 bp corresponding to housekeeping genes *glcK* and *thrS* (Fig. 3.3.1). The purified PCR products were sequenced and an online BLAST search confirmed the presence of the *Streptococcus thermophilus* *glcK* and *thrS* genes with 99% homology with reported sequences.
Fig.3.3.1 Gel electrophoresis (1% agarose) of PCR amplification products of genes \textit{glcK} and \textit{thrS}. Lanes A and D: GeneRuler® 100bp DNA Ladder (Thermo, CA); Lane B: Amplification products from \textit{S. thermophilus} ST21 genomic DNA with primer pair targeting housekeeping gene \textit{glcK}; Lane C: Amplification products from \textit{S. thermophilus} ST21 genomic DNA with primer pair targeting housekeeping gene \textit{thrS}. 
3.3.2 Antibiotic sensitivity and transformability of *Streptococcus thermophilus* ST21

After electroporation, cells transformed with pGKV 210 as well as the wild type *S. thermophilus* ST21 were recovered, then plated on selective medium. Erythromycin-resistant transformants were isolated following electroporation while no colonies were obtained from the negative control. This confirmed that the wild type *Streptococcus thermophilus* ST21 strain did not possess erythromycin resistance but the Em\(^r\) gene was acquired from foreign DNA by the transformant.

Since the transformability of the strain *S. thermophilus* ST21 had not been previously reported, and to ensure the strain is transformable, we used the plasmid pGKV 210, a well documented rolling-circle replication plasmid and an *E. coli/LAB* shuttle vector, which was reported to successfully transform *Streptococcus thermophilus* (Su et al. 2002). As a result, *S. thermophilus* ST21 was successfully transformed with plasmid pGKV 210 as demonstrated by the isolation of erythromycin-resistant clones. Random single colonies were picked and sub-cultured prior to plasmid extraction and restriction enzyme digestion (*Figure 3.3.2*).

After restriction enzyme digestion, plasmid pGKV 210 has a size around 4.4 kb. The plasmid extracted from *E. coli* showed 3 bands on agarose gel with sizes about 3 kb, 4.3 kb, and 9 kb, which indicated the presence of the three forms of plasmid DNA: supercoiled, linear, and circular plasmid, respectively. After digestion with *Bam*HI, supercoiled and circular plasmids were linearized at the *Bam*HI restrict site, thus there was only one band at around 4.3 kb (*Figure 3.3.2a*), which is the expected size of plasmid pGKV 210.
The electropherogram of plasmids extracted from *Streptococcus thermophilus* ST21 transformed with pGKV 210 and wild type *S. thermophilus* ST21 (as a negative control) showed that *S. thermophilus* ST21 harboured endogenous plasmids with sizes from 1.5 kb to around 5 kb (Fig 3.3.2b). Therefore, the plasmid profile of pGKV 210 *S. thermophilus* transformants showed the presence of transformed plasmid bands with the expected sizes of around 3 kb, 4.3 kb, and 9 kb as well as endogenous plasmids with sizes of 1.5 kb, 2 kb and 5 kb. The lack of the 4.4 kb band in the wild type *Streptococcus thermophilus* ST21 also provided evidence that the plasmid DNA associated with the erythromycin resistance was successfully transformed. It has been reported that a few *Streptococcus thermophilus* strains can harbor more than one plasmid without interfering with normal cell function (Somkuti and Steinberg 1986). In this study, the introduction of the foreign plasmid DNA did not cause endogenous plasmid(s) to be lost according to the electrophoresis result. Compared to the plasmid profile of wild type ST21, the pGKV 210 transformed cells not only harbors foreign plasmid pGKV 210, but also the complete set of the *S. thermophilus* ST21 endogenous plasmid(s) as the wild type. Thus, in this study, it can be concluded that, on one hand, even though harboring endogenous plasmid(s), *Streptococcus thermophilus* ST21 was successfully transformed with the shuttle vector pGKV 210 carrying the erythromycin resistance gene; on the other hand, the cryptic endogenous plasmid(s) in *S. thermophilus* ST21 and pGKV 210 were compatible for no plasmid loss was observed after the host cells were transformed with the heterologous plasmid.
Fig. 3.3.2 Gel electrophoresis (1% agarose) of plasmid pGKV 210 following restriction enzyme digestion. (a) Lane A: λHind III DNA Marker (Thermo Fisher Scientific); Lane B, C: plasmids extracted from E. coli harboring pGKV 210 digested with BamHI; Lane D: plasmid extracted from E. coli without enzyme digestion. (b) Lane A and H: 1kb DNA ladder (New England biolabs); Lane B and D: plasmids extracted from Mutant #1 and #2 digested with BamHI, respectively; Lane C and E: plasmids extracted from Mutant #1, #2 without enzyme digestion, respectively; Lane F: plasmids extracted from wildtype Streptococcus thermophilus ST21 digested with BamHI; Lane G: plasmid from wildtype Streptococcus thermophilus ST21 without enzyme digestion.
Transformation efficiency (number of transformants / µg plasmid DNA) was calculated based on the ratio of the amount of plasmid DNA (measured by NanoDrop 1000) and the numbers of antibiotic-resistant clones obtained on the selective medium (Table 3.3.2).

**Table 3.3.2.** Transformation Efficiency of pGKV 210 into *S. thermophilus* ST21

<table>
<thead>
<tr>
<th>Hours of cell recovery/antibiotics resistance gene expression</th>
<th>Transformation efficiency (Number of Transformants / µg plasmid DNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>((0.22 \pm 0.03) \times 10^3)</td>
</tr>
<tr>
<td>3</td>
<td>((0.32 \pm 0.14) \times 10^3)</td>
</tr>
<tr>
<td>6</td>
<td>((0.60 \pm 0.23) \times 10^3)</td>
</tr>
<tr>
<td>9</td>
<td>((1.23 \pm 0.26) \times 10^3)</td>
</tr>
<tr>
<td>12</td>
<td>((1.42 \pm 0.43) \times 10^3)</td>
</tr>
</tbody>
</table>

* Values are the means of three samples ± standard deviation

One-way ANOVA showed that post-electroporation cell recovery has a significant effect on the transformation efficiency \((p=0.0005)\). Groups with significantly different means were identified using Tukey's test. As a result, only the 6 h post-electroporation recovery treatment group and the 9 h post-electroporation recovery treatment group showed significant difference \((p=0.036)\) meaning transformation efficiency was significantly higher after cells were recovered in an antibiotic-free medium at 37°C for more than 6 hours, but did not increase after 9-hour incubation time.

The shuttle vector pGKV 210 (4.4 kb) was derived from the *L. lactis* plasmid pWV01. As engineered, plasmid pGKV 210 encodes an Em<sup>r</sup> gene as a selection marker. It is reported that to obtain erythromycin resistance, modification after gene translation is needed, so
the time for the newly transformed bacteria to recover is the key for getting higher transformation efficiency (Horinouchi and Weisblum, 1980). As reported, the transformation efficiency of pGKV 210 in plasmid-free *Streptococcus thermophilus* could reach around $7 \times 10^3$ transformants per µg plasmid DNA compared to $6 \times 10^3$ transformants in *E. coli*. However, the transformation efficiency may vary from strain to strain. Considering there is no significant difference after 9 h post-electroporation recovery time, to obtain satisfying transformation efficiency, optimal cell recovery/erythromycin resistance gene expression is expected 6 h to 9 h post-electroporation and these conditions were adopted in the later experiments.

In conclusion, the results showed that although *Streptococcus thermophilus* ST21 harbors endogenous plasmid(s), it is still transformable, and the transformants have the ability to harbor the foreign plasmid DNA (at least up to 4 kb) and its own plasmid(s) simultaneously.

3.3.3 Vector plasmid pIL252 can be maintained and replicated in *Streptococcus thermophilus* ST21

As mentioned above, plasmid pGKV 210 follows rolling-circle replication (also known as the sigma mode of replication). However, rolling-circle replication (RCR) has its limitation for carrying large-insert DNA. It is reported that the plasmid segregational instability increases with the size of DNA inserts in RCR plasmids (Shareck et al. 2004). To overcome this segregational instability, and stably maintain a large foreign DNA fragment, a theta-replication plasmid would be preferred. Thus, the use of plasmid pIL252 was investigated.
Plasmid pIL252 is a low copy number plasmid vector derived from pAMβ-1. It was originally designed for molecular cloning in *Lactococcus lactis* (Simon and Chopin 1988). Although *S. thermophilus* is genetically close to *L. lactis*, and it has been reported that pIL252 can successfully transform *Streptococcus pneumoniae*, yet few reports have been published about the transformation of *S. thermophilus* with plasmid pIL252. Moreover, it has been reported that not all *S. thermophilus* strains can be transformed with plasmid pIL253, a higher copy number version of pIL252, which was also derived from pAMβ-1 (O'Sullivan and Fitzgerald 1999). Therefore, before transforming pEDSB (pIL252m + 20 kb EPA/DHA gene cluster), it was necessary to test the compatibility between *S. thermophilus* ST21 and the vector plasmid.

Two different methods of transformation, electroporation and protoplast-curing, were used to transform *S. thermophilus* with plasmid pIL252. Erythromycin-resistant colonies were obtained following electroporation. However, no transformants were obtained when the polyethylene glycol (PEG)-mediated protoplast transformation was used. Transformation efficiency of electroporation was calculated and an average of 45 erythromycin-resistant transformants per µg plasmid DNA was obtained under the electroporation setting of a single pulse of 2,500 V lasting 5 ms, capacitance of 25 µF and resistance of 200 Ω. Figure. 3.3.3 shows the electrophoretograms of the plasmid DNA isolated from the *E. coli* strain harboring pIL252, *Streptococcus thermophilus* ST21 transformed with pIL252 and wild type *S. thermophilus* ST21 following restriction enzyme digestion. As mentioned, plasmid pIL252 is a low copy number plasmid. Thus, the bands on the gel are of lower intensity compared to those from pGKV210 transformants. Nevertheless, the expected bands could be identified. Compared to wild
type *S. thermophilus* ST21, before *Bam*HI digestion, the pIL252-transformed *S. thermophilus* cells exhibited extra bands with sizes around 9 kb and 4.5 kb. After restriction enzyme digestion, the 9 kb band (circular plasmid) disappeared and was replaced by a more intense 4 kb band (linear plasmid). *S. thermophilus* ST21 can harbor more than one plasmid, but wild type *S. thermophilus* ST21 is susceptible to erythromycin. Thus, the survival of the *S. thermophilus* transformants carrying pIL252 under erythromycin pressure proved that the low copy number vector pIL252 can replicate and express the erythromycin resistance gene in recombinant *Streptococcus thermophilus*.
Fig. 3.3.3. Gel electrophoretogram (1% agarose) of plasmid pIL252 following restriction enzyme digestion. Extracted plasmid from *L. lactis* IL1403<sup>+</sup>, *S. thermophilus* transformants and wildtype *S. thermophilus* ST21 were digested by restriction enzyme *BamHI*. Lanes A and H: 1 kb DNA ladder (New England Biolabs); Lane B: pIL252 extracted from *L. lactis* IL1403<sup>+</sup> digested with *BamHI*; Lane C: undigested pIL252 from *L. lactis* IL1403<sup>+</sup>; Lane D: plasmid from *S. thermophilus* pIL252 transformant digested with *BamHI*; Lane E: undigested plasmid from *S. thermophilus* pIL252 transformant; Lane F: wildtype ST21 plasmid digested with *BamHI*; Lane G: undigested wildtype ST21 plasmid.
3.3.4 Transformation of plasmid pEDSB (pIL252m + 20 kb EPA/DHA gene cluster) into *Streptococcus thermophilus* ST21

Based on the fact that the plasmid pIL252 can be maintained and replicate in *Streptococcus thermophilus* ST21, an attempt was made to transform *S. thermophilus* ST21 with the 20 kb EPA/DHA gene cluster.

Plasmid pEDSB was constructed previously by Amiri-Jami et al. (2014) and harbored by *Lactococcus lactis* subsp. *cremoris* MGED20. It was constructed based on the ligation of the 20 kb gene cluster to a modified pIL252 vector. The original 140-bp multiple cloning site of pIL252 was replaced by a 160-bp synthetic sequence containing a *Not*I restriction site (Amiri-Jami, Lapointe, and Griffiths 2014). Considering the similarity between *L. lactis* and *S. thermophilus*, it was hypothesized that *S. thermophilus* ST21 could also be transformed with pEDSB. In order to obtain recombinant transformants, electroporation and PEG-mediated protoplast transformation were used. As a result, 5 erythromycin resistant transformants were obtained following electroporation, but no transformants were obtained from the protoplast transformation.

The result of colony PCR (Fig. 3.3.4.1) showed that specific band (around 1.8 kb) was able to be amplified by the oligonucleotide gene *pfaA* primer pair in pEDSB transformed *S. thermophilus* no.1 (Lane C); whereas single bands with smaller molecular weight (around 1.1 kb) were observed in pEDSB transformed *S. thermophilus* strain no.3, no.4 and no.5 indicating gene *pfaA* in these strains might have been modified by their host. *Shewanella baltica* MAC1 was used as a positive control (Lane B).
Fig. 3.3.4.1. Gel electrophoretogram of amplicons following colony PCR targeting the \textit{pfaA} gene. Lanes A and H: GeneRuler 100 bp DNA Ladder (Life Technologies, Burlington, ON, Canada); lane B, colony PCR product from \textit{Lactococcus lactis} subsp. \textit{cremoris} MGED20 harboring pEDSB; lanes C, D, E, F and G: colony PCR products from pEDSB transformed \textit{Streptococcus thermophilus} no.1, no.2, no.3, no.4 and no.5, respectively.
However, to confirm the accurate insert size of the EPA/DHA synthesizing gene cluster in the pfaA-positive mutants, NotI HF restriction enzyme digestion was necessary.

Plasmid extracted from *Lactococcus lactis* subsp. cremoris MG1363 harboring pEDSB and its NotI-digested product was set as a positive control (Figure 3.3.4.2a). As expected, NotI HF was able to cut the plasmid pEDSB at the NotI site into 2 fragments to produce plasmid pIL252m (4.5 kb) and the 20 kb EPA/DHA gene cluster. As for *S. thermophilus* transformants, after NotI HF digestion, only mutants no.1, no.3 and no.4 showed inserts around 20 kb and vectors around 4 kb. The restriction enzyme digestion patterns of the pEDSB-transformed *S. thermophilus* are shown in Figure 3.3.4.2b.
Fig. 3.3.4.2a. Electrophoretograms of plasmids extracted from *Lactococcus lactis* subsp. *cremoris* MG1363 and digested with *NotI* HF. Lanes A and F, λHind III DNA Marker (Thermo Fisher Scientific); lanes B and D, plasmids extracted from *L. lactis* subsp. *cremoris* MGED20 digested with *NotI* HF; land C and E, plasmid extracted from *L. lactis* subsp. *cremoris* MGED20 without digestion.
Fig. 3.3.4.2b. Electrophoretograms obtained following restriction enzyme (NotI HF) digestion of plasmids extracted from pIL252m + 20kb inserted mutants. Lanes A and H: λHindIII DNA Markers; lanes B, D, F: plasmid extracted from transformants no.1, no.3, and no.4 digested by NotI HF, respectively; lanes C, E, F: plasmid extracted from transformants no.1, no.3, and no.4, respectively.
Although it has been reported that *Streptococcus thermophilus* is transformable, to date, few reports have described the cloning of a large DNA fragment into *S. thermophilus*. However, in this study, a low copy number theta-replicating vector plasmid carrying a 20 kb gene cluster insert was successfully transformed into *S. thermophilus* ST21. However, the transformation efficiency of the recombinant plasmid pEDSB (pIL252m + 20 kb EPA/DHA gene cluster) was very low because of its size. On average, only 2.5 transformants were obtained from 1 µg plasmid DNA (plasmid pEDSB).

Colony PCR indicated that four of the erythromycin-resistant isolates, colonies no.1, no.3, no.4 and no. 5 showed a single band following amplification of gene *pfaA*, which is an essential gene for the biosynthesis of EPA and DHA (Amiri-Jami et al. 2014). Furthermore, just like the positive control, transformant no.1 gave an expected band around 1.8 kb, indicating that it may possess the entire *pfaA* gene; whereas the others (colonies no.3, no.4 and no.5) might harbor a partial sequence of the *pfaA* gene.

Moreover, *NotI* HF digestion of the isolated plasmid DNA from the clones (no.1, no.3 and no.4) resulted in two bands of the correct size, 20 kb and 4.5 kb associated with the EPA/DHA gene cluster inset and the vector plasmid (pIL252m), respectively. Also, compared to the plasmids extracted from recombinant *L. lactis* harboring pEDSB (**Fig. 3.3.4.2a** lane B,C) and the wild type *S. thermophilus* ST21 harboring its endogenous plasmid (**Fig. 3.3.4.2b** lane F,G), the gel-profile of the pEDSB transformed *S. thermophilus* ST21 showed the capacity for harboring both the transformed recombinant plasmid pEDSB and the endogenous plasmid(s). As seen in **Fig. 3.3.4.2b**, smaller endogenous plasmids and the larger foreign plasmid pEDSB (24.5 kb) were all maintained by the *S. thermophilus* transformants. The recombinant *S. thermophilus*
clones carrying plasmid pEDSB were named as *Streptococcus thermophilus* STED20 no.1, no.3 and no.4.

### 3.3.5 Gas Chromatography (GC) and Gas Chromatography - Mass Spectrometry (GC-MS) analysis of the pEDSB transformed *Streptococcus thermophilus* clones

The fatty acid methyl esters derived from *S. thermophilus* STED20 no.1, no.3 and no.4 were analyzed by gas chromatography. *Shewanella baltica* MAC1 and wild type *Streptococcus thermophilus* ST21 were set as positive and negative controls, respectively. Results showed that the three clones revealed peaks near the retention time obtained with the EPA standard, but no peak was observed at the retention time associated with the DHA standard (Fig. 3.3.5.1).

Gas chromatography – mass spectrometry was used to confirm the identity of the putative EPA peak. With gas chromatography - mass spectrometry, compounds in the samples were first separated by their retention time and then, molecules were ionized (and thus charged) by bombarding with a high-energy electron beam. Molecular ions and fragment ions would be accelerated by manipulation of the charged particles through the mass spectrometer. Ions traveled down the path based on their mass to charge ratio (m/z). The mass analyzer sorts the ions according to m/z and the detector records the abundance of each m/z. Thus, different compounds generate different “unique” mass spectra as “fingerprints”.

Mass spectrometry confirmed that the putative EPA peak produced by *S. thermophilus* STED20 no.1 was cis-5,8,11,14,17-eicosapentaenoic acid, while no EPA was detected by GC-MS in the negative control (Fig. 3.3.5.2).
The yield of recombinant EPA produced by *S. thermophilus* STED20 no.1 was also calculated from the peak areas of the known concentration of standard EPA and the EPA peak area observed for the transformant. At 25°C, *S. thermophilus* STED20 produced 0.27 ± 0.07 mg EPA per gram dry cell weigh (mean value ± standard deviation), while at 15°C, the positive control, wild type *Shewanella baltica* MAC1, produced 1.27 ± 0.41 mg EPA per gram dry cell weight.
Fig. 3.3.5.1 GC results (a) Gas chromatogram of standard EPA and DHA. The retention time of the EPA and DHA standards are 27.924 min and 31.329 min, respectively (b) Gas chromatogram of total fatty acid methyl esters prepared from *S. thermophilus* STED20 no.1. Fatty acids were extracted from cells grown in 1% (w/v) GM17 broth supplemented with 2.5 µg ml⁻¹ erythromycin for 4 days at 25°C without shaking.
**Fig. 3.3.5.2 GC-MS result (a)** The comparison of electron impact mass spectra of the putative EPA peak in *S. thermophilus* STED20 no.1 (red) with that of the known standard mass spectra for EPA standard (blue). **(b)** Results given by the GC-MS suggested that the structure of the compound corresponding to the putative EPA peak of *S. thermophilus* STED20 no.1 was cis-5,8,11,14,17-eicosapentaenoic acid.
3.3.6 Stability of recombinant plasmid pEDSB in *Streptococcus thermophilus* STED20

To test the stability of the plasmid pEDSB in *Streptococcus thermophilus* STED20, an overnight culture grown at 37°C was sub-cultured and grown under the same conditions overnight. However, analysis using restriction enzyme digestion, PCR of the *pfaA* gene and GC and GC-MS all indicated that the complete recombinant plasmid was not stably maintained in *S. thermophilus* STED20, after an additional overnight culture.

Although the recombinant strain *S. thermophilus* STED20 was confirmed to have the ability to harbor recombinant plasmid pEDSB and thus produce recombinant EPA in the first generation, the recombinant plasmid cannot be maintained stably in the transgenic *S. thermophilus*. **Fig. 3.3.6** illustrates the restriction enzyme digestion patterns obtained from the sub-cultured pEDSB-transformed *S. thermophilus*. A band of size 4.5 kb was observed, but no bands were seen at 20 kb after *NotI* HF digestion. Instead, bands around 8 kb, which did not exist before in any of the erythromycin resistant transformants, were observed after *NotI* digestion. Furthermore, none of the transformants possessed the *pfaA* gene (no amplified bands were shown after PCR amplification), which resulted in no EPA being detected by gas chromatography.

In addition, the transformants could not hold the insert at -80°C after repetitive thawing and refreezing. After transformation by electroporation, all the erythromycin positive transformants were grown overnight under antibiotic pressure and immediately stored at -80°C in 15% sterile glycerol. However, similar results to those above following *NotI* digestion and colony PCR were obtained using the revived *Streptococcus thermophilus* STED20 cells after a month; indicating that, faced with more stressful extreme conditions
when being stored (-80°C), the recombinant plasmid pEDSB in the transgenic *Streptococcus thermophilus* STED20 lost the unnecessary inserted 20 kb EPA/DHA gene cluster to adapt to the harsh environment.

![Gel electrophoretogram following restriction enzyme NotI HF digestion of plasmids extracted from the re-subcultured *S. thermophilus* STED20. Lanes A and H: λHindIII DNA Markers; Lanes B, D, F: plasmid extracted from re-subcultured *S. thermophilus* STED20 no.1, no.3, and no.4 digested by *NotI* HF, respectively; Lanes C, E, F: plasmid extracted from re-subcultured *S. thermophilus* STED20 no.1, no.3, and no.4, respectively.](image-url)
3.4 General discussion

In the past two decades, numerous genetically engineered vector plasmids for lactic acid bacteria were constructed. To genetically modify *S. thermophilus*, different plasmid vectors were proposed and engineered for different applications. The first constructed vectors, pA2 and pA33, were based on *S. thermophilus* endogenous rolling cycle replicon, however, they were later found to suffer from segregational instability (Mercenier et al. 1988). Later on, more and more plasmid vectors with higher transformation efficiency were engineered. For example, the pMEU series *E. coli/S. thermophilus* shuttle vectors based on pER8 (RC replicon) and ori of *E. coli* showed high efficiency for *S. thermophilus* and *E. coli* transformation with high segregational and structural stability (Solaiman & Somkuti 1993); large conjugative plasmid vectors pIP501, pAMβ1 and their derivatives (including pIL252 used in this research) based on theta mode of replication exhibited high segregational and structural stability even with large heterologous DNA inserts (Kiewiet et al. 1993). Thus, to obtain the appropriate transformation efficiency and plasmid stability, it is essential to use the proper type of plasmid.

It has been documented that the plasmid replication mode has an important impact on the host-range, stability, and copy number of the plasmids (Shareck et al. 2004). Generally speaking, sigma and theta modes of replication are the most common replication mechanisms in LAB.

Sigma-replicating plasmids are also known as rolling-circle replicating (RCR) plasmids. It is assumed that sigma-replicating plasmids are the most widespread plasmids among Gram-positive bacteria. Normally, they have high-copy number in their host cells. However, due to the generation and accumulation of single-stranded DNA, RCR
plasmids usually have low segregational stability (Vujcic and Topisirovic, 1993). Also, even though successful case of cloning DNA fragments up to 10 kb using RCR plasmid pLS69 has been reported, the stability of the plasmid was only limited in homologous host (but not heterologous host) (Stassi et al. 1981). On the contrary, it was generally reported that a large DNA insertion significantly increased the segregational instability of RCR plasmid vectors (Fernández-López et al. 2015; Leer et al. 1992). Theta-replicating plasmids are found mostly at low copy number in LAB. Unlike RCR plasmids, theta-replicating plasmids have higher segregational stability even when containing a large DNA insert. Plus, theta-replicating mechanism derived vectors tend to suffer less from plasmid incompatibility problems (Kiewiet et al. 1993).

In this chapter, we first had to test the transformability of the target bacterium *S. thermophilus* ST21, as there was no previous publication documenting the transformation efficiency of this strain. To transform *S. thermophilus* ST21, and to optimize conditions for transformation, a well-documented rolling-circle replication broad host-range shuttle vector pGKV 210 (derived from pWV01) was used (van der Vossen, Kok, and Venema 1985). As a result, an acceptable transformation rate for *S. thermophilus* ST21 of about $10^3$ transformants per µg plasmid DNA was achieved. At least 6 h of post-electroporation cell recovery time was necessary for the transformants to express the erythromycin resistance gene.

However, due to the instability of pGKV 210 (an RCR plasmid), as mentioned above, theta-replicating plasmids with low copy number and ability to stably hold a large foreign DNA insert were preferred to express the 20 kb EPA/DHA gene cluster. Plasmid pIL252 is a low copy number vector plasmid derived from theta-replicating plasmid pAMβ-1.
Originally, it was designed for molecular cloning in *Lactococcus lactis* (Simon and Chopin 1988).

*Streptococcus thermophilus* ST21 was successfully transformed with vector plasmid pIL252, but at a relatively low transformation rate (45 transformants per µg plasmid DNA). It is noticeable that *Streptococcus thermophilus* ST21 harbors endogenous plasmid(s). As one of the most important extrachromosomal DNA elements, plasmid DNA widely exists among different bacterial species. One of the most important reasons that bacteria can colonize and adapt to different environments in a relatively short period of time is their ability to acquire novel traits from extrachromosomal DNA elements from their ecological niches. These foreign DNA elements might encode functional genes such as antibiotic resistance, specific enzymes of metabolic pathways, or virulence genes that may give the host cells survival advantages. Once accepted by the new host, such new genetic traits will help the new species become dominant in the population, thus, the co-existing relationship between plasmids and hosts becomes heritable and after generations, endogenous. In this sense, bacterial plasmids played a critical role in bacterial evolution.

In this study, based on the stability and compatibility of the endogenous plasmid(s), the cryptic plasmid(s) may be essential in the bacterial life cycle, nutrient metabolism, or environment adaptation of *S. thermophilus* ST21.

As mentioned above, recombinant plasmid carrying the 20 kb EPA/DHA gene cluster pEDSB was constructed based on modified pIL252. The original 140-bp multiple cloning cite of pIL252 was replaced by a 160-bp synthetic sequence containing a *Not*I restriction site (Amiri-Jami et al, 2014). As *S. thermophilus* ST21 can be transformed by pIL252 and
thus gained erythromycin resistance, it was hypothesized that *S. thermophilus* ST21 can maintain plasmid pEDSB.

However, the transformation rate of pEDSB was lower than pIL252. Five erythromycin resistant transformants were obtained. Fortunately, four of them were positive following PCR targeting the *pfaA* gene, indicating that they possessed the *pfaA* gene (completely or partially), even though the band was of the expected size for only one of the transformants. Among these four *pfaA*-positive clones, clones no.1, no.3 and no.4 harbored a DNA fragment with the correct size (20 kb) expected for the EPA/DHA insert.

GC and GC-MS results confirmed that *S. thermophilus* transformed with pEDSB resulted in the production of EPA in *Streptococcus thermophilus* STED20 no.1 (0.27 mg EPA g\(^{-1}\) CDW in 1% GM17 broth supplemented with erythromycin at 25°C), whereas wild type *Streptococcus thermophilus* normally produces unsaturated fatty acids only up to C20:1 (eicosenoic acid) (Beal, Fonseca, and Corrieu 2001). Also, recombinant *S. thermophilus* STED20 no.1 produced more EPA than recombinant *L. lactis* MGED20 (0.12 ± 0.04 mg EPA g\(^{-1}\) CDW) (Amiri-Jami et al. 2014). On the other hand, it was reported that the pEDSB transformed *L. lactis* subsp. *cremoris* MG1363 was able to produce recombinant DHA (Amiri-Jami et al. 2014); however, no recombinant production of DHA was detected by *Streptococcus thermophilus* STED20. This suggests that *Streptococcus thermophilus* might not have the ability to correctly translate/modify the *pfaE* gene, which is responsible for the production of bacterial DHA (Okuyama et al. 2007).

Additionally, the quantity of recombinant EPA produced by *S. thermophilus* ST21 harboring pEDSB was still lower than for recombinant *E. coli* harboring recombinant plasmid pfBS-PS, which is possibly the result of the low copy number of the vector.
plasmid pIL252 in *Streptococcus thermophilus*. Also, as a Gram-positive bacterium, *Streptococcus thermophilus* only has an inner cell membrane; thus, the absence of an outer membrane may directly lead to a reduced amount of EPA being produced or incorporated into membranes.

To date, there is no report describing the successful cloning of any large size (as large as 20 kb) heterogeneous gene(s) in *Streptococcus thermophilus*. The successful transformation of the 20 kb EPA/DHA gene cluster as well as the expression of recombinant EPA in transformants has great potential. The successful cloning of the gene cluster was credited mostly to the design and construction of the plasmid pEDSB (pIL252m + 20 kb EPA/DHA gene cluster). It is reported that pIL252 can stably maintain a large size DNA insert and replicate in some *L. lactis, Lactobacillus* and *Streptococcus* strains (Simon and Chopin 1988; Domingues et al. 2013; Amiri-Jami, Lapointe, and Griffiths 2014). Most importantly, the test of vector-host compatibility in this chapter proved that pIL252 could replicate in *Streptococcus thermophilus* ST21. As reported by Amiri-Jami et al. (2014), attempts to clone the 20 kb EPA/DHA to *L. lactis* subsp. *cremoris* MG1363 were all unsuccessful except for the one using pIL252m as cloning vector, indicating that plasmid pIL252 carries a replicon that can be replicated by LAB strains. However, it is noteworthy that even though pIL252 was reported to be a cloning vector that can stably carry a large foreign DNA insertion, the recombinant plasmid pEDSB was not stable in *Streptococcus thermophilus* STED20. The 20 kb EPA/DHA gene cluster was lost following the second sub-culture. Furthermore, plasmid loss also happened in the stock cultures, which were stored at -80°C supplemented with 15% sterile glycerol, as no 20 kb insert was observed in the revived cells.
It is well documented that recombinant plasmid stability relies on both the recombinant plasmid and the host cells. Normally, plasmid loss is caused by segregational instability and structural instability (Smith and Bidochka 1998). The inserted foreign DNA fragment might encode protein/peptides that interrupt normal cellular functions. Also, it is reported that the different G+C content or codon usage preference between the original organisms and the secondary hosts might inhibit protein expression, thus leading to plasmid instability (Terpe 2006). A new host with similar base composition and codon usage/frequency is optimal. In this case, the inserted 20 kb heterogeneous gene cluster might have caused a great metabolic burden to *S. thermophilus* or have undesirable high-frequent codons that are rarely used by *S. thermophiles*, whereas it did not happen in *L. lactis* or *E. coli*. Thus, transformants with such plasmids were overtaken by those plasmid-negative mutants due to the survival disadvantage of the former.

The instability of plasmid pEDSB can also be explained by the CRISPR/Cas system in *Streptococcus thermophilus*. The existence of CRISPR/Cas system (clustered regularly interspaced short palindromic repeats/CRISPR associated), a bacterial “adaptive immune system”, in *Streptococcus thermophilus* has been well documented (Mojica et al. 2005). “CRISPR” are a family of DNA sequences with repetitive palindromic sequences separated by regularly sized, non-repetitive intervening spacer DNA sequences (normally originating from phages or plasmids) in genomes of some bacteria and Archaea (Sapranauskas et al. 2011).

To cope with the crisis of being exposed to foreign genetic elements such as phages and plasmids, bacteria and archaea have evolved an elaborate adaptive immune system to degrade harmful/unnecessary foreign nucleic acid (Garneau et al. 2010; Labrie, Samson,
and Moineau 2010). Previously, it was reported that the CRISPR1/Cas system in *Streptococcus thermophilus* could provide acquired immunity against phages and self-replicating plasmids with antibiotic-resistance marker by integrating novel spacer DNA sequences accordingly in the invading foreign genetic elements, which resulted in the specific cleavage of these invasive DNA sequences (Garneau et al. 2010). Garneau et al. (2010) found that after a certain number of generations, *S. thermophilus* transformed with plasmid encoding a chloramphenicol-resistance gene became sensitive to chloramphenicol again. As *S. thermophilus* obtained immunity to “invasive” plasmids, sequence analysis of CRISPR1 in the plasmid-lost-colonies was able to identify spacers highly homologous to the former inserted plasmid DNA. Thus, it was concluded that the CRISPR1/Cas system also helps bacteria defeat the “invasion” of plasmids, the foreign, virus-like DNA element.

The mechanism of CRISPR/Cas mediated adaptive immunity can be described as follows. Firstly, CRISPR loci integrate short DNA sequences as spacer DNA sequences from invading genetic elements. The integrated repeat-spacers will then be transcribed into small CRISPR RNAs. As guides, these CRISPR RNAs can help the Cas proteins to target the invasive DNA. Finally, the functional domains carried by the Cas proteins, such as nucleases, will result in cleavage or interference of foreign DNA sequences (Garneau et al. 2010).

As plasmid pEDSB carries a large size 20 kb foreign DNA fragment encoding 5 *pfa* genes originating from a Gram-negative bacterium, *Shewanella baltica* MAC1, the transformation of such recombinant plasmid into the Gram-positive bacterium, *Streptococcus thermophilus* may cause disordered cellular function. In this case,
**Streptococcus thermophilus** might have activated the CRISPR/Cas system to modify this foreign DNA element to protect itself. However, as the vector plasmid pIL252 encodes an erythromycin-resistance gene, which offers an obvious survival advantage, the mutated *Streptococcus thermophilus* STED20 with only the vector plasmid pIL252 survived under erythromycin selection pressure.

In conclusion, a recombinant *Streptococcus thermophilus* strain harboring the 20 kb EPA/DHA biosynthesis gene cluster was successfully constructed. As a result, recombinant production of EPA was detected in *S. thermophilus* STED20. However, the yield of recombinant EPA produced by *S. thermophilus* STED20 was relatively low compared to EPA produced by transformed *E. coli*; and no recombinant DHA could be detected. The low production of EPA could be due to the lack of an outer cell membrane in *S. thermophilus*, while the reason of the absence of recombinant DHA might be failure of proper translation or modification of the *pfaE* gene. Furthermore, the inserted 20 kb EPA/DHA gene cluster could not be stably maintained in *S. thermophilus* STED20, which may be due to 1) the low copy number of the vector plasmid pIL252 in *S. thermophilus*, 2) the large size of the EPA/DHA synthesis gene cluster (20 kb), 3) the extra metabolic burden from the protein encoded by the heterogeneous genes, 4) the variant preference of codons, base composition between the secondary host, *S. thermophilus*, and the original host, *Shewanella baltica*, and/or 5) the existence of CRISPR/Cas adaptive immune system in *S. thermophilus*.

Therefore, to stably maintain the EPA/DHA gene cluster, integrating the foreign DNA fragment into the chromosome of *S. thermophilus* using integrating plasmid vector such as pGEM5- and pUC19-derived plasmids can be considered in the future. However,
considering the large insertion of heterogeneous gene cluster in chromosomal DNA might cause severe cellular dysfunction, a different food-grade bacterial strain may be necessary for stable production of omega-3 fatty acids. *Lactobacillus acidophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus* in this case, might be possible secondary hosts of this marine-derived 20 kb EPA/DHA biosynthesis gene cluster. First of all, *Lactobacillus acidophilus* is a very commonly used bacteria in commercialized probiotic products; whereas *Lactobacillus delbrueckii* subsp. *bulgaricus* is used in yogurt production. Moreover, in the past two decades, many successful cases of transforming *Lactobacillus acidophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus* with various vector plasmids, such as pNZ123 (rolling-circle-replicating) to transform *L. acidophilus* and pNZ12 (rolling-circle-replicating) or pGB305Δ (theta-replicating) to transform *L. delbrueckii* subsp. *bulgaricus*, have been well documented (Kim et al. 2005; Serror et al. 2002).
4 General discussion and future directions

4.1 Introduction
It is widely accepted that omega-3 fatty acids, especially EPA and DHA are beneficial to human health. So far, cold-water marine fish are still the principal dietary source of EPA and DHA (Robles Medina et al. 1998). However, due to limitations such as the instability of the fish supply, water pollution, and the diminishing fish resources, alternative dietary EPA and DHA sources are urgently needed to fill the expanding fish oil supplement market. Due to the high efficiency and relatively low cost of microbial fermentations, genetically engineered microorganisms are widely used in the pharmaceutical industry to produce bioactive compounds such as IL-2, INF-α, recombinant insulin, etc. Recently, microalgae were used for industrial production of EPA and DHA (Arterburn et al. 2007). However, the production of EPA/DHA from mass-cultured microalgae is costly and inefficient (Chi et al. 2009). Thus, to meet the market demand of dietary EPA and DHA, a novel dietary source of these acids which can be produced with higher efficiency and lower cost is needed.

Microorganisms were once thought to lack the series of enzymes needed to synthesize PUFAs (Erwin et al. 1964). However, because of the great improvement in analytical instruments, especially gas chromatography (GC) and gas chromatography-mass spectrometry (GC-MS), the detection of trace components in a complex bio-matrix has become feasible, which has led to the identification of bacterial species of marine origin that have the ability to produce very long-chain PUFAs such as EPA and/or DHA (Hirota 2005; Bowman and Gosink 1998; Amiri-Jami et al. 2006). In the last two decades, countless effort has been spent on isolating effective EPA/DHA genes found in marine
bacteria so that they may be cloned into other bacteria (Metz 2001; Nishida et al. 2006; Okuyama et al. 2007).

The first milestone of EPA/DHA gene mining was the successful cloning of a 38 kb EPA-producing gene cluster from *Shewanella putrefaciens* SCRC-2378 to *E. coli* (Yazawa 1996). Recently, Amiri-Jami and Griffiths (2010) found a 35 kb gene cluster in *Shewanella baltica* MAC1 responsible for producing EPA/DHA, which was subsequently successfully cloned into different *E. coli* strains. Moreover, heterogeneous expression of this gene cluster in *E. coli* strains resulted in significantly higher production of EPA and DHA at a wider temperature range compared to the wild type *Shewanella baltica* MAC1 (Amiri-Jami and Griffiths 2010). More recently, the 35 kb gene cluster was re-sized to 20 kb by deleting the predicted genes that were unnecessary for EPA/DHA production (Amiri-Jami et al. 2014). This 20 kb EPA/DHA gene cluster resulted in the production of even higher levels of EPA and DHA in transformed *E. coli* strains. Furthermore, this gene cluster has been successfully cloned into *L. lactis* subsp. *cremoris* MG1363, a food grade microorganism widely used in the dairy industry, using a modified vector plasmid pIL252m (Amiri-Jami et al. 2014).

Based on this knowledge, to better apply the EPA/DHA biosynthesis genes in the food industry in the future, it was decided to test the stability of the recombinant plasmid pfBS-PS (8 kb pCC1 FOS + 20 kb EPA/DHA gene cluster) in *E. coli* EPI300T1 transformants without the supplementation of chloramphenicol to provide selective pressure. Furthermore, to further reveal the true potential of the gene cluster, an attempt was made to clone the 20 kb EPA/DHA gene cluster into *S. thermophilus* ST21, to see if
it can be stably maintained in this GRAS dairy starter and lead to the stable production of recombinant EPA and DHA.

4.2 General discussion

As mentioned, marine fish as the present primary dietary source of EPA and DHA has its limitations, therefore, scientists have turned their attention to developing novel, high efficiency, low cost, edible, and tasty alternative sources to fill the big gap in the market. With the development of microbiology, molecular biology and biotechnology, microorganisms, as a relatively “simple” organism (compared to plants and animals), have attracted the most attention. Up to now, although already commercialized, limitations of mass-production of EPA and DHA using microalgae (high cost, low efficiency, difficult to refine and separate EPA and DHA from the complex product of fermentation) still need to be solved. However, the discovery of high-efficiency EPA/DHA biosynthesis gene cluster(s) using the PKS pathway from marine bacteria and the successful cloning and expression of these genes to other microorganisms, especially *L. lactis* subsp. *cremoris*, have opened a new window to bring other food-grade microorganisms as “candidates” to be potential alternative dietary EPA and DHA sources.

4.2.1 Plasmid stability

In the first part of this study, the stability of the recombinant plasmid harboring the 20 kb EPA/DHA producing gene cluster was tested in the recombinant *E. coli* EPI300T1 cultivated for 20 consecutive days (with periotic sub-culture every 4 days) in the absence of chloramphenicol. Previous studies revealed that EPA and/or DHA biosynthesis gene clusters could be transformed to *E. coli* strains and lead to the production of EPA and/or DHA under antibiotic selection pressure (Yazawa 1996; Orikasa et al. 2006; Amiri-Jami
Nevertheless, few data were reported regarding the long-term stability of these genes in recombinants without the supplementation of antibiotics during consecutive growth cycles. In this study, the 20 kb EPA/DHA gene cluster carried by pCC1 FOS showed good stability in *E. coli* EPI300T1; on the other hand, the same gene cluster carried by pIL252 suffered from plasmid instability in *S. thermophilus*.

In the food industry, the stability of recombinant plasmids without antibiotic selection pressure during culture is a primary concern. During fermentation, plasmid loss is the main reason for the production of a non-active product; leading to considerable financial losses. Plus, according to Food and Drug Regulations – C.R.C., c.870 (Section B.01.048.1 N and C.01.606), antibiotics are strictly regulated in the dairy industry and animal production. Under the regulation, no antibiotic was allowed to be added during dairy production. Furthermore all raw milk must be proved to be antibiotic-free before entering processing plants. Considering the importance of antibiotics for plasmid stability, it is necessary to assess the long-term stability of the recombinants in the absence of antibiotics to ensure economic profit and compliance with legal requirements.

For the recombinant stability test in *E. coli*, to better simulate an industrial fermentation, the stability of the pfBS-PS without any antibiotic supplementation in the medium was studied for 20 days with periodic sub-culture every 4 days. As a result, even without adding antibiotic, the recombinant plasmid exhibited good stability. Plasmid extraction followed by restriction enzyme digestion at the end of the cultivation clearly showed the existence of the 20 kb EPA/DHA inserted DNA fragment and the 4 kb vector plasmid, which indicated that the transformed *E. coli* strains can stably maintain the recombinant plasmid during prolonged cultivation in the presence or absence of chloramphenicol.
Plasmids are inheritable, extra-chromosomal DNA elements that are present in prokaryotes and in some lower eukaryotes. Plasmids have the ability to shuttle from species to species and thus mediate exchanges of genetic material and genetic traits (Andrup et al. 2003). Thus, unlike chromosomal DNA, properties obtained by plasmid transformation are not as stable due to the different vector-host compatibility between its original host and secondary host. Also, as plasmids are inherited independently from the bacterial chromosome, specific phenotypes given by plasmids are not normally essential for the growth of bacteria. Only plasmids that introduce obvious survival advantages to the host cells can be stably maintained and passed by the transformants (e.g. production of bacteriocins, exopolysaccharides; resistance to antibiotics or heavy metals). It has been reported that cultivation conditions, especially presence of antibiotics, strongly influence the stability of transformed plasmids (Gerdes et al. 1986; Rohde 1995). The ability to stably maintain foreign DNA also depends on the characteristics and compatibility of the host cells and the recombinant DNA (Smith & Bidochka 1998). As heterologous genes, the foreign DNA element in the new host may encode unnecessary transcription product/protein, which can cause a substantial metabolic burden to the transformants and thus put them in a growth/survival-disadvantaged position. Plus, the base composition and codon encoded by the foreign DNA element might be rarely used (or cannot be recognized by the messenger RNA) to the new host, and thus lead to translation failure. Hence, transformants carrying foreign DNA fragments will be replaced by the plasmid/DNA insert-free mutants, which have growth/survival advantages. Generally speaking, larger inserted DNA sequences are associated with more complex transcription products and thus require greater energy expenditure by cells. Also, the copy number of
the vector plasmid may affect the stability of the recombinant plasmid. Normally, high copy number plasmids can ensure the amount of plasmid DNA that passes to the daughter cells. However, high copy number plasmids also correspond to high replication and transcription metabolism. Therefore, with large inserted DNA fragments, high copy number plasmid vectors can cause plasmid segregational instability (Friehs 2003).

In this study, the absence of chloramphenicol and long-term cultivation did not cause plasmid instability. The first possible explanation may be that the 20 kb EPA/DHA insert did not cause extra metabolic burden to *E. coli* EPI300T1 transformants. *Shewanella baltica* and *E. coli* are both gram-negative bacteria. Also, genomic analysis of *Shewanella* genes encoding enzymes of EPA synthesis showed that, out of the 11 predicted domains, one of the FAS domains was homologous to the *E. coli* FabA dehydrase (Metz et al. 2001), which indicated that the *E. coli* fatty acid synthesis pathway and the *Shewanella* EPA synthesis pathway share common enzyme(s). Thus, the genetic similarity might explain the excellent compatibility and stability of the heterogenous EPA/DHA gene cluster in recombinant *E. coli* strains.

Also, the genetically engineered host strain and plasmid vector combination, *E. coli* EPI300T1 plasmid pCC FOS1, allowed us to control the plasmid copy number to obtain better plasmid stability. *E. coli* EPI300T1 encodes the *trfA* gene, which can activate the high-copy replication origin *oriV* in pCC FOS1, and *trfA* is regulated by the $P_{araBAD}$ ($P_{BAD}$) promoter and AraC protein. Without inducing the expression of $P_{BAD}$ promoter, gene *trfA* would not be activated, and plasmid pCC FOS1 would maintain a very low copy number in *E. coli* EPI300T1. Although one domain is predicted to be highly similar to the *E. coli* *FabA* dehydrase, the EPA/DHA gene cluster is still large and heterogeneous to *E. coli*. So,
to avoid causing a high metabolic burden to the transformants, in this study, gene \textit{trfA} was inactivated during cultivation, so that the recombinant plasmid pfBS-PS stayed at a low copy number. Thus, the deficiency of mass replication of the recombinant plasmid pfBS-PS may also be a key to the long-term stability in the production of bacterial EPA and DHA in transformed \textit{E. coli} strains.

The pAM\(\beta\)1-based plasmid pIL252 was chosen as a vector plasmid to carry the 20 kb gene cluster and to transform \textit{S. thermophilus} ST21. First and foremost, as a lactococcal theta-replicating plasmid, pIL252 can maintain larger foreign DNA inserts more stably (Shareck et al. 2004). Secondly, pIL252 was successfully transformed into certain streptococcal strains as a low copy number plasmid (Domingues et al. 2013). Moreover, the recombinant plasmid pEDSB using the modified pIL252 as vector to carry the 20 kb EPA/DHA gene cluster has been constructed, cloned and expressed in \textit{L. lactis} subsp. \textit{cremoris} MG1363, which is a LAB strain similar to \textit{S. thermophilus}.

The compatibility test between host cells and empty plasmid vector showed that the vector plasmid pIL252 could replicate in \textit{S. thermophilus} ST21 and express the erythromycin-resistance gene even with its endogenous plasmid(s). Also, the low concentrations of DNA, as evidenced by spectrophotometry and the faint bands obtained following gel electrophoresis of the pIL252 transformants, suggested that plasmid pIL252 has a low copy number in \textit{S. thermophilus} ST21.

However, when \textit{S. thermophilus} ST21 was transformed with pIL252 carrying the 20 kb EPA/DHA gene cluster, the recombinant plasmid could not be maintained even with one more overnight sub-culture.
To date, very few successful studies describing transformation of large size foreign DNA sequences to *Streptococcus thermophilus* have been reported. Furthermore, so far, the recombinant strain *S. thermophilus* STED20 is the only streptococcal strain reported that can produce omega-3 fatty acids. Thus, in association with the long-term recombinant plasmid stability in *E. coli* without antibiotic selective pressure, the transformation of the EPA/DHA biosynthesis gene cluster from *Shewanella baltica* MAC1 into *S. thermophilus* has great industrial potential if problems related to the stability of this EPA/DHA gene cluster in *S. thermophilus* can be resolved.

Apart from the reasons mentioned above, in *S. thermophilus*, the existence of CRISPR/Cas might also be a main reason for the plasmid instability. Recently, it has been found that *Streptococcus thermophilus* has developed an elaborate bacterial “adaptive immune response” called CRISPR/Cas system to protect them from invasion by foreign DNA elements when exposed to phages and plasmids. In the chromosome of *S. thermophilus*, a family of DNA sequences with repetitive palindromic sequences separated by intervening spacer DNA sequences (highly homologues to phage and plasmid sequences) exist; this family of DNA was named CRISPR (Mojica et al. 2005). To protect itself, when exposed to foreign genetic elements, *S. thermophilus* will integrate novel spacer DNA from the invasive DNA fragments into its chromosome to generate a CRISPR sequence. Once the integrated spacers are transcribed to small CRISPR RNAs, the CRISPR RNAs will work as a guide to help the Cas proteins to locate the homologous gene sequences in the invasive DNA. As a result, the functional domain, such as nuclease, in Cas proteins will cleave/modify the hazardous DNA element to remove the threat (Garneau et al. 2010).
The 20 kb EPA/DHA gene cluster is a relatively large and complicated heterogeneous DNA element to \textit{S. thermophilus}. The \textit{pfa} genes from \textit{Shewanella} might cause an extremely heavy metabolic burden in \textit{S. thermophilus}. Thus, it is reasonable to hypothesize that it might have activated its CRISPR/Cas system to protect itself.

4.2.2 Recombinant production of EPA and DHA

The 20 kb gene cluster used in our previous study was reported to produce 7-fold more EPA (in \textit{E. coli} grown in LB broth supplemented with chloramphenicol) than the wild type \textit{Shewanella baltica} MAC1 (in marine broth) cultured at 15°C for 3 days (Amiri-Jami & Griffiths 2010).

Gas chromatography confirmed that the plasmid pfBS-PS was successfully expressed in \textit{E. coli}. The average yield of recombinant EPA from the last generation of recombinant \textit{E. coli} grown in chloramphenicol-free LB broth was 71.27 mg g\textsuperscript{-1} CDW. It showed no statistical significant difference (P >0.05) compared to the EPA yield provided by the first generation (81.70 mg g\textsuperscript{-1} CDW). Two-way ANOVA was also used to test the influence of antibiotic supplementation and cultivation time on the synthesis of bacterial omega-3 fatty acids. Neither the cultivation duration nor the supplementation of chloramphenicol significantly affected the yield of recombinant EPA and DHA. Analysis using GC and GC-MS showed that \textit{S. thermophilus} STED20 no.1 was able to synthesis EPA (1.27 mg EPA g\textsuperscript{-1} CDW in 1% GM17 broth supplemented with erythromycin at 25°C), while the wild type \textit{Streptococcus thermophilus} was only able to synthesize unsaturated fatty acids with chain lengths up to C20:1 (eicosenoic acid) (Beal et al. 2001). However, no recombinant DHA could be detected.
Compared to the average production of recombinant EPA/DHA in *E. coli* EPI300T1 reported in this study, the yield of EPA given by *S. thermophilus* STED20 was significantly lower. Above all, the plasmid instability may be the most important reason for the low production of recombinant n-3 fatty acids. As mentioned, *S. thermophilus* is reported to produce PUFA up to eicosenoic acid; therefore, the bio-synthesis of EPA and DHA in *S. thermophilus* is all based on the proper expression of the gene cluster. Secondly, as a Gram-positive bacterium, *S. thermophilus* only has an inner cell membrane, thus, it is reasonable to conclude that the lack of an outer membrane is an important factor leading to the low production of recombinant EPA/DHA. In addition, similarly it was reported that the amount of EPA and DHA produced by genetically modified *L. lactis* subsp. *cremoris* MG1363 was also lower than that observed for the recombinant *E. coli* strains (Amiri-Jami et al. 2014). Moreover, the low copy number of the vector plasmid pIL252 in *S. thermophilus* may also be responsible for the low EPA/DHA production as it directly leads to a low rate of transcription and expression of the EPA/DHA gene cluster. Furthermore, unlike *E. coli*, which shares a homologous gene *fabA* to the 20 kb EPA/DHA gene cluster, none of the genes in the EPA/DHA gene cluster were predicted similar to any of the genes encoded by *S. thermophilus*. It is suggested that *S. thermophilus* might have a completely different pathway for fatty acid metabolism from *Shewanella*. In this case, the heterogeneous EPA/DHA gene cluster may have to compete with the streptococcal endogenous fatty acid synthesis genes, and thus leads to a significantly lower yield of EPA. Also, the two PUFA synthesis pathways might also explain the absence of recombinant DHA in *S. thermophilus* STED20, as the DHA synthesized through PKS pathway might have been further metabolized by *S.*
thermophilus and redirected into its own fatty acid metabolic system, which might have converted DHA to other fatty acids. It is reported that the phosphopantetheinyl transferase encoded by gene pfaE and the 3-ketoacyl synthase encoded by pfaB are responsible for synthesizing DHA (Orikasa et al. 2006; Orikasa et al. 2009). The lack of DHA production might also suggest that S. thermophilus cannot correctly transcribe pfaE to generate the bioactive and correct conformation of the phosphopantetheinyl transferase and 3-ketoacyl synthase.

In conclusion, we first confirmed the long-term stability of the recombinant plasmid pfBS-PS (pCC FOS1 + 20 kb EPA/DHA gene cluster) and long-term stability of the production of recombinant EPA/DHA in E. coli EPI300T1 without the application of antibiotic selective pressure. Secondly, we successfully constructed a recombinant strain, S. thermophilus STED20, harboring recombinant plasmid pEDSB (pIL252m + 20 kb EPA/DHA biosynthesis gene cluster), which resulted in the production of recombinant EPA. However, due to the low production of recombinant omega-3 fatty acids provided by S. thermophilus STED20 and the recombinant plasmid instability, further research is needed.

4.3 Future directions

1) To stably transform the 20 kb EPA/DHA gene cluster into S. thermophilus ST21, integration vectors can be used to integrate the gene cluster into its chromosomal DNA.

2) To apply the 20 kb EPA/DHA gene cluster to LAB strains, other genetically stable dairy starter cultures sharing close phylogenetic relationship with S. thermophilus and L.
lactis sup. cremoris, such as Lactobacillus and Bifidobacterium, should be considered as a secondary host.

3) Instead of prokaryotic cells, lower eukaryotic cell expression system (model organisms), such as yeast and Arabidopsis thaliana, can also be considered as a secondary host.

4) To further discover the EPA/DHA biosynthesis mechanism given by the 20 kb gene cluster and to outline the PKS pathway regulation, real-time PCR should be performed to measure the expression level of all 5 of the pfa genes.

5) To increase the expression level of 20 kb gene cluster, expression systems (e.g. Baculovirus expression vector system) with strong promoters (e.g. Baculovirus 39K protein promoter and Baculovirus basic protein promoter), can be genetically engineered.


Arterburn, L M et al. 2007. Bioequivalence of docosahexaenoic acid from different algal oils in capsules and in a DHA-fortified food. Lipids.


**Bolotin, A et al. 2004.** Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nature biotechnology*.

**Bowman, J P & Gosink, J J. 1998.** *Colwellia demingiae* sp. nov., *Colwellia hornerae* sp. nov., *Colwellia rossensis* sp. nov. and *Colwellia psychrotropa* sp. nov.: psychrophilic Antarctic species with the ability to synthesize docosahexaenoic acid (22:6n-3). *International journal of systematic bacteriology* doi:10.1099/00207713-48-4-1171.


**Burdge, G C. 1998.** The role of docosahexaenoic acid in brain development and fetal alcohol syndrome.. *Biochemical Society transactions* 26 (2) (maiatzak): 246–252.


Caughey, G E et al. 1996. The effect on human tumor necrosis factor alpha and interleukin 1 beta production of diets enriched in n-3 fatty acids from vegetable oil or fish oil.. The American journal of clinical nutrition.


Chiu, Chih-Chiang et al. 2012. Associations between n–3 PUFA concentrations and cognitive function after recovery from late-life depression. The American journal of clinical nutrition.


Ezez. 2011. Safety of Probiotics to Reduce Risk and Prevent or Treat Disease (apirolak 1).


Garneau, Josiane E et al. 2010. The CRISPR/Cas bacterial immune system cleaves


**Gerdes, K et al. 1986.** Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the *E. coli* relB operon.. *The EMBO journal* 5 (8) (abuztuak): 2023–2029.


**H J Busscher et al. 1997.** *Streptococcus thermophilus* and its biosurfactants inhibit adhesion by *Candida* spp. on silicone rubber.. *Applied and environmental microbiology* 63 (10) (urriak 1): 3810.


**Hare, Peter D & Chua, Nam-Hai. 2002.** Excision of selectable marker genes from transgenic plants. *Nature Biotechnology* 20 (6) (ekainak 1). doi:10.1038/nbt0602-575.


**Helland, Ingrid B et al. 2008.** Effect of supplementing pregnant and lactating mothers with n-3 very-long-chain fatty acids on children's IQ and body mass index at 7 years of age.. *Pediatrics* 122 (2) (abuztuak): e472–9. doi:10.1542/peds.2007-2762.

**Heude, Barbara; Ducimetière, Pierre & Berr, Claudine. 2003.** Cognitive decline and fatty acid composition of erythrocyte membranes—The EVA Study. *The American


Leer, R J et al. 1992. Structural and functional analysis of two cryptic plasmids from Lactobacillus pentosus MD353 and Lactobacillus plantarum ATCC 8014. Molecular and General ...


Morita, N; Tanaka, M & Okuyama, H. 2000. Biosynthesis of fatty acids in the
docosahexaenoic acid-producing bacterium *Moritella marina* strain MP-1.


**Olsen, Sjurdur F et al. 2008.** Fish oil intake compared with olive oil intake in late pregnancy and asthma in the offspring: 16 y of registry-based follow-up from a randomized controlled trial.. *The American journal of clinical nutrition* 88 (1) (uztailak): 167–175.


**Quint, Joseph F & Fulco, Armand J. 1973.** The Biosynthesis of Unsaturated Fatty
Acids by Bacilli. *Journal of Biological Chemistry.*


Simon, Daniel & Chopin, Alain. 1988b. Construction of a vector plasmid family and


**Snow, A A; Andow, D A & Gepts, P. 2005.** Genetically engineered organisms and the environment: current status and recommendations 1. *Ecological Applications*.


**Su, Ping et al. 2002.** Cloning vectors for *Streptococcus thermophilus* derived from a native plasmid.. *FEMS microbiology letters* 216 (1) (urriak 29): 43–47.


