# The Development of a Thermodynamic Model for Antisense RNA Design and an Electro-transformation Protocol to Introduce Auxotrophic Genes for Enhancing Eicosapentaenoic Acid Yield from *Pythium irregulare*

Yang Yue

Thesis submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
In
Biological Systems Engineering

Ryan Senger, Chair
Chenming Zhang, Committee Member
David Bevan, Committee Member

December 7, 2011

Blacksburg, VA

Keywords: eicosapentaenoic acid(EPA), fermentation and analysis, antisense RNA, thermodynamic model

# The Development of a Thermodynamic Model for Antisense RNA Design and an Electro-transformation Protocol to Introduce Auxotrophic Genes for Enhancing Eicosapentaenoic acid Yield from *Pythium irregulare*

Yang Yue

ABSTRACT

Eicosapentaenoic acid (EPA, C20:5, n-3) is a long chain crucial unsaturated fatty acid, essential for the regulation of critical biological functions in humans. Its benefits include the therapeutic treatment of cardiovascular disease, schizophrenia and Alzheimer's disease. The fungus *Pythium irregulare* (ATCC 10951) has great potential as a natural EPA producer. In this study, the electroporation conditions for *P. irregulare* were determined. The auxotrophic selectable genes *ura*, *trp* and *his* were respectively cloned into the plasmid pESC to construct shuttle vectors. Electroporation with 2.0kV and a 0.2cm cuvette was applied as the most effective condition for heterogeneous genes transformation. The yield and content of EPA and other components of total fatty acids (TFA) were further determined by the FAME approach with GC, as well as the analysis of biomass. The EPA content in *P. irregulare* with heterologous pESC-TRP vector reached 16.68 mg/g if cultured in auxotrophic medium, which showed a 52.33% increase compared to the wild-type *P. irregulare*. The maximum of EPA yield was 98.52 mg/L from *P. irregulare* containing the pESC-URA plasmid, a 32.28% increase over the wild-type. However, the maximum cell dried weight of these two

organisms were respectively 6.13g/L and 5.3g/L, significantly less than the 6.80g/L of the wild-type. Not only was a feasible approach detected to electro-transform and increase the EPA yield of *P. irregulare*, this study also inferred that ω-6 route was mainly involved in the EPA biosynthesis in this organism.

An antisense RNA (asRNA) thermodynamic model was developed to design new asRNA constructs capable of fine-tuning gene expression knockdown. The asRNA technology is now identified as an effective and specific method for regulating microbial gene expression at the posttranscriptional level. This is done by targeting mRNA molecules. Although the study of regulation by small RNAs is advanced in eukaryotes, the regulation of expression through artificially introducing antisense oligodeoxynucleotides into host is still being developed in prokaryotes. To study the thermodynamics of asRNA and mRNA binding, (i) the fluorescence protein genes *GFP* and *mCherry* were separately cloned into the common pUC19 vector and (ii) antisense GFP and antisense mCherry DNA fragments were randomly amplified and inserted into the constructed plasmid under the control of an additional *plac* promoter and terminator. The expression level of fluorescence reporter proteins was determined by plate reader in this combinatorial study. A thermodynamic model to describe the relationship between asRNA binding and observed expression level was created. The study indicates two factors that minimum binding energy of the asRNA-mRNA complex and the percentage of asRNA binding mRNA were crucial for regulating the expression level. The correlation relationship between gene expression level and binding percentage multiplied by the minimum binding energy was found to show a good correlation between the thermodynamic parameters and the observed level of gene expression. The model has the

potential to predict the sequence of asRNA and the approach will ultimately be applied to cyanobacteria to increase lipids production. Here, the long-term approach is to build metabolic switches from asRNA that can turn “on/off” various cellular programs and metabolic pathways at will in a fine-tuned manner. This will allow engineers to control metabolic activity in response to reactor conditions.

# ACKNOWLEDGEMENTS

I would like to take this opportunity to be grateful to my advisor Dr. Ryan Senger for his guidance and support throughout this research. His scientific attitude and passion for research influences and motives me all my life.

I would also like to thank my committee members, Dr. Chenming Zhang and Dr. David Bevan for your kind guidance and support.

Also, I would like say thank you to my lab mates. It is my honor to work with you all and learn from you all.

Last but not least, I would like to thank my family, my parents and my wife. Without your support, I could not be able to step this far. Your love will be with me forever.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

## 1. Overview and motivation

### 1.1. EPA production from *Pythium irregulare*

Eicosapentaenoic acid (EPA, C20:5, n-3) is a long chain unsaturated fatty acid, essential for the regulation of critical biological functions in humans. It is also beneficial for the therapeutic treatment on cardiovascular disease, schizophrenia and Alzheimer's disease [1]. The traditional source of EPA is extraction from sea fish oil. This approach provides most of the current EPA supply on the market. However, the objectionable taste and odor of fish oil in addition to concerns over mercury contamination as well as the fluctuating supply have inhibited the application of EPA in the food industry [2]. In addition to the EPA recovered from fish oil, it is also derived from algae and fungi, which are also primary producers of omega-3 fatty acids. Significant research has begun to focus on yielding EPA from fungi fermentation and optimizing EPA production through genetic manipulations and metabolic engineering of regulatory pathways. To implement this strategy in the laboratory of Dr. Zhiyou Wen at Virginia Tech, *Pythium irregulare* was chosen as the host for metabolic engineering experiments to optimize the production of EPA. The advantages of *P. irregulare* include its high growth rates on economical substrates and great potential as a natural EPA producer. Efforts in this research to metabolically engineer *P. irregulare* were initiated by attempting to electro-transform and replicate a plasmid in this organism. The electroporation conditions were studied fist, and are described in Chapter 2. This work was performed under the guidance of Dr. Wen in the Biological Systems Engineering Department at Virginia Tech. The impact of electro-transformation was then studied on the EPA and total fatty acids (TFA) yield. The influence of introducing heterologous auxotrophic vectors into *P. irregulare* was also determined on EPA and TFA yield and organism content in this research. The influence of culture conditions and medium components on the growth of *P. irregulare* was also studied and their impacts on EPA yield were determined. These

experiments were designed to ultimately increase EPA content of the organism through metabolic engineering. The EPA biosynthesis pathway is commonly controlled by Δ5 desaturase, Δ6 desaturase, and Δ6 elongases [3]. Regulation of the expression levels of the genes coding these crucial enzymes is a potential target for metabolic engineering approaches to alter metabolic flux and increase EPA biosynthesis.

### 1.2. Fine-tuning gene expression with antisense technology

Antisense RNA (asRNA) technology is now identified as an effective method for regulating microbial gene expression at the translational level. The approach is to introduce (into a host cell) a recombinant expression vector that contains an asRNA open reading frame, which is controlled by a promoter of choice. Once the asRNA is transcribed, it specifically inhibits the synthesis of a target protein by hybridizing with a targeted mRNA molecule, from which the target protein is translated. This process has been found in many cases to reduce the specific expression level of a targeted gene. One example involves the transformation of asRNA into *Clostridium acetobutylicum* to improve the fermentation of acetone and butanol [4]. In this case, asRNA was used as a substitute for a gene knockout. However, current metabolic engineering strategies being developed in the laboratory of Dr. Ryan Senger in the Biological Systems Engineering Department at Virginia Tech require partial and "fine-tuned" gene knockdowns (e.g., 30% reduction in gene expression). The asRNA approach is the most feasible way of accomplishing this task. However, the design of asRNA for fine-tuned gene expression knockdown has not yet been accomplished.

Although the study of regulation by antisense technologies is advanced in eukaryotes, the regulation of expression through artificially introducing asRNA into prokaryotes is still being

developed. The main problem lies on the delivery, suppression efficiency, stability, and the danger of off-target effects such as the silencing of unwanted genes. Therefore, the improvement of asRNA research on the above problems could provide a feasible and promising method on gene expression level regulation. The research of this study is focused on developing a new thermodynamic model to enable asRNA design. The model is expected to predict the quantitative gene expression level (relative to an un-inhibited control) given the primary sequence input of the asRNA. To study the thermodynamics of asRNA and mRNA hybridization, the fluorescent protein reporter genes for green fluorescent protein (*GFP*) and *mCherry* were cloned separately into the common pUC19 vector. The asRNA fragments designed for *GFP* and *mCherry* were obtained from a combinatorial PCR design and were inserted into the constructed plasmid under the control of an additional inducible *plac* promoter and terminator. The expression levels of fluorescent reporter protein were determined in real time by plate reader in this study. A thermodynamic model to describe the relationship between the particular asRNA used and the observed expression level of the fluorescent reporter was developed. This work was done under the guidance of Dr. Senger at Virginia Tech. This study and results are fully described in Chapter 3 of this thesis.

### 1.3. The overall focus of metabolic engineering of this thesis

Both Chapters 2 and 3 of this thesis describe the development of tools related to metabolic engineering. An electro-transformation protocol is required to insert new genes to engineer *P. irregulare* metabolism for increased EPA production. The development of asRNA technologies will enable further "fine-tuned" manipulations of metabolism to ultimately improve EPA yield. The thermodynamic asRNA model will be applied to predict the regulatory effects of a designed asRNA sequence. With this technology, the long-term approach is to build metabolic switches

from asRNA that can turn "on/off" various cellular programs and metabolic pathways at will, allowing engineers to control metabolic activity in response to reactor conditions.

**1.4. References:**


1. Simopoulos, A. P. Essential fatty acids in health and chronic disease. Am. J. Clin. Nutr. 1999, 70, 560S–569S.
2. Innis, S. M. Fatty acids and early human development. Early Hum. DeV. 2007, 83, 761–766
3. Astrid Meyer, Helene Kirsch, Frédéric Domergue. Novel fatty acid elongases and their use for the reconstitution of docosahexaenoic acid biosynthesis.The Journal of Lipid Research, October 2004, 45, 1899-1909
4. Seshu B. Tummala, Stefan G. Junne, and Eleftherios T. Papoutsakis. Antisense RNA Downregulation of Coenzyme A Transferase Combined with Alcohol-Aldehyde Dehydrogenase Overexpression Leads to Predominantly Alcohologenic Clostridium acetobutylicum Fermentations. J. Bacteriol. 2003, 185(12):3644

## 2. The effects of electro-transformation and autotrophic selection on the growth and eicosapentaenoic acid yield from *Pythium irregulare*

### 2.1. Introduction and Background

#### 2.1.1. Omega-3 Polyunsaturated Fatty Acids

Omega-3 polyunsaturated fatty acids are unsaturated fatty acids that contain a double bond occurring after the third central carbon atom from the methyl end. Alpha-linolenic acid (ALA, 18:3), eicosapentaenoic acid (EPA, 20:5), and docosahexaenoic acid (DHA, 22:6) are significant members of the omega-3 polyunsaturated fatty acids family. These have important functions related to human health and are regarded as a necessary dietary supplement. They regulate cholesterol transport and fatty acid biosynthesis [1]. Despite the fact that the omega-3 polyunsaturated acids play such important physiological roles, diet is the most effective and feasible way to obtain them [2].

#### 2.1.2. Eicosapentaenoic acid

Eicosapentaenoic acid (EPA, C20:5, n-3) is a long chain crucial unsaturated fatty acid which is essential for the regulation of critical biological functions in humans. Health benefits also include the therapeutic treatment of cardiovascular disease, cancers, schizophrenia and Alzheimer's disease [3, 4]. EPA is also a precursor to docosahexaenoic acid (DHA) which has beneficial effects on infant brain and retinal development [5]. As an important member of the omega-3 fatty acid family, EPA like other omega-3 polyunsaturated fatty acids, can be synthesized from alpha linolenic acid (ALA) in the human body, but this process is much less efficient than obtaining it from food. Therefore, EPA is a commonly added nutrient in several foods, especially those designed for babies [5].

### 2.1.3. The structure and biosynthesis of EPA

The chemical structure of EPA is show in Fig. 2.1. As other polyunsaturated fatty acids, EPA is biosynthesized from acetate via extension and desaturase of a short chain fatty acid. Oleic acid (18:1, ω-9) and ALA (18:3, ω-3) are two essential intermediates in the synthesis process. Fig. 2.2 shows the biosynthesis steps of EPA from acetate.

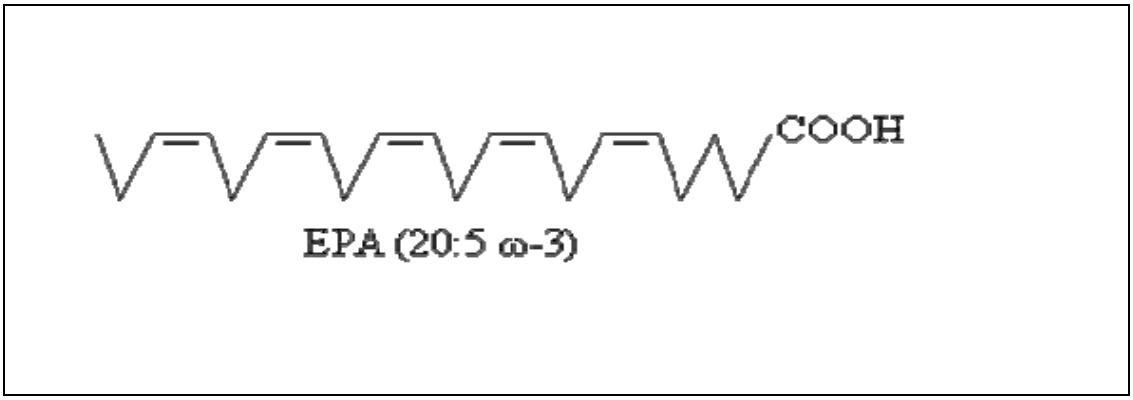



Fig. 2.1.The chemical structure of eicosapentaenoic acid (EPA).

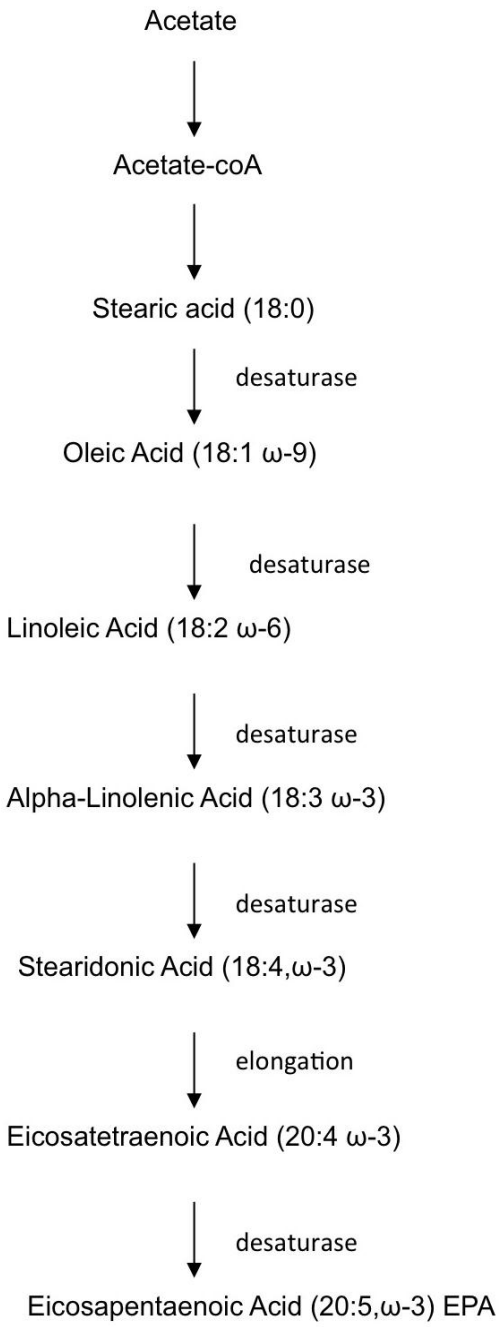



Fig. 2.2. Biosynthetic pathway of EPA from acetate.

### 2.1.4. EPA sources

Extracting polyunsaturated fatty acids from fatty fish oil such as anchovies, salmon, and sardines is the traditional route to produce EPA. This method limits the application and yield because of (i) the undesirable taste and odor, (ii) heavy metal contamination (especially mercury) [6], (iii) potential for overfishing, and (iv) cost of the product [7]. Flaxseed, canola oil, and walnuts are also sources for unsaturated fatty acids, but these contain little EPA and adequate alpha-linolenic acid (ALA) [8].Therefore, a novel and sustainable source to yield EPA is beneficial and sought highly. This research is inspired by fact that fish do not produce EPA naturally. Rather, they obtain it from the algae they consume [9].The fungus (also classified as an algae) *Pythium irregulare* was studied as an alternative source to produce EPA using crude glycerol as the carbon source. Crude glycerol is abundant and a major byproduct of the biodiesel industry. It is economical, and its availability is adequate for EPA production. Previous research of the Wen Lab at Virginia Tech indicated *P. irregulare* shows great potential as a natural EPA producer, with a yield of 14 mg/L following optimization of culture conditions. The growth medium contained 30 g/L of crude glycerol and 10 g/L yeast extract [10]. Using biodiesel-derived crude glycerol as a carbon source for EPA production, microalgae *P. irregulare* solves two problems. First, it provides a method of disposal for waste glycerol. Second, it offers a sustainable source of EPA.

### 2.1.5. P. irregulare as an EPA producer

*P. irregulare* is found worldwide and is a soil borne plant pathogen. This organism causes blight and root rot of plants and has been isolated from hundreds of plants [11]. Once classified as a fungus, *P. irregulare* has recently been re-classified as an algae [12]. It reproduces by both sexual and asexual mechanisms, and spores are also formed by the mature mycelia. Glucose can

be used as the sole carbon source to produce long chain unsaturated fatty acids [13]. *P. irregulare* can also utilize waste flux or sweet whey with an increasing yield of EPA [14] at 12 ℃. The medium pH has been found to play an important role in EPA production. The optimized pH was found to be between 6 and 7, while significant inhibition was observed at pH 8 [15]. Both temperature and pH have been found to control the mycelia morphology of clumps or pellets [16]. A previous study in the Wen Lab showed the mycelia morphology is also affected by the rotation speed of culture agitation. The Wen lab used crude glycerol as a less expensive substrate for the fermentation of *P. irregulare* to produce EPA with the control of strain growth in small pellets. The biomass yield reached 9.50 g/L at 18 ℃.

### 2.1.6. Focus of this research

From the metabolic pathway of EPA synthesis, Δ5 desaturase, Δ6 desaturase, and Δ6 elongases are commonly considered the key enzymes for the biosynthesis of EPA. A general idea to optimize the EPA yield is to overexpress the genes coding these enzymes in the host. However, for *P. irregulare*, the transformation conditions required for the heterologous genes were still unknown as well as the exact EPA synthesis route. To obtain a high EPA yielding engineered strain, the electroporation conditions were determined first. Then, the influence of culture conditions and medium components on the growth of *P. irregulare* and EPA yield was also explored in this research.

## 2.2. Materials and Methods

### 2.2.1. Strain, growth medium, and culture conditions

The organism *P. irregulare* (ATCC 10951) was used in this research to determine effective electroporation conditions for DNA transfer and the impacts on EPA yield. *P. irregulare* was cultured in medium containng 30 g/L glucose and 10 g/L yeast extract (control medium). Although previous research results showed crude glycerol from the biodiesel industry could be applied as an economical carbon source that also increased EPA yield [10], glucose was used as the sole carbon source in this study to ensure the stability and replication of growth by the culture. The pH of the medium was adjusted to 6.8 before autoclaving at 121 ℃ for 15 min. *P. irregulare* was cultured and incubated into 50 mL liquid medium in a 250 mL flask. After 5-7 days of incubation at 26 ℃ in an orbital shaker set to 170 rpm, the harvested mycelia were disrupted by agitation in distilled water containing glass beads. This suspension contained spores and was saved to serve as a future inoculum.

The control medium and trial medium were both prepared for electroporation and EPA production experiments. For the control medium, the components mentioned above were used for culturing wild-type *P. irregular*. The trial medium contained glucose , yeast nitrogen base, and *Saccharomyces cerevisiae* complete synthetic defined supplement deficient in a particular component (tryptophan, histidine, or uracil). Glucose was utilized as the sole carbon source with the final at a concentration of 30 g/L for both medium. For the nitrogen source, the control medium contained of 10g/L yeast exact, while *Saccharomyces cerevisiae* complete synthetic defined supplement deficient of uracil (URA), histidine (HIS), or tryptophan (TRP) were added separately for the three different trial medium. These media were prepared and autoclaved with

the concentrations of 0.77 g/L, 0.77 g/L, and 0.72 g/L for the URA, HIS, and TRP deficient media, respectively. Yeast nitrogen base, without amino acids, was as also added to the trial media to a concentration of 6.7 g/L after filter sterilization. The trial media were used for determining the survival of host after the electroporation of auxotrophic selective heterologous vectors. As a crucial amino acid was deficient in each of the trial media, only the positive colonies with vectors were able to survive. The culture was inoculated at 2% (v/v) of the total volume into 50 mL of liquid medium and incubated at 26 ℃ with a shaker speed of 170 rpm. The cultures were grown for 7 days in all experiments.

### 2.2.2. Construction of plasmids

The plasmids applied in this study were offered from the lab of Dr. Bingyu Zhao at Virginia Tech. They were constructed from pESC yeast epitope tagging vectors (Invitrogen). These vectors all provided (i) the origin of replication units in *E. coli* and yeast *S. cerevisiae*., (ii) ampicillin resistance, (iii) two MCS sites regulated separately promoters *Gal1* and *Gal10* in opposing directions. An auxotrophic selectable marker gene (*HIS3, TRP1, or URA3*) was respectively located in the plasmid backbone to select and maintain the expression vector in host cells. The structure of the *HIS3* auxotrophic plasmid is shown in Fig. 2.3.

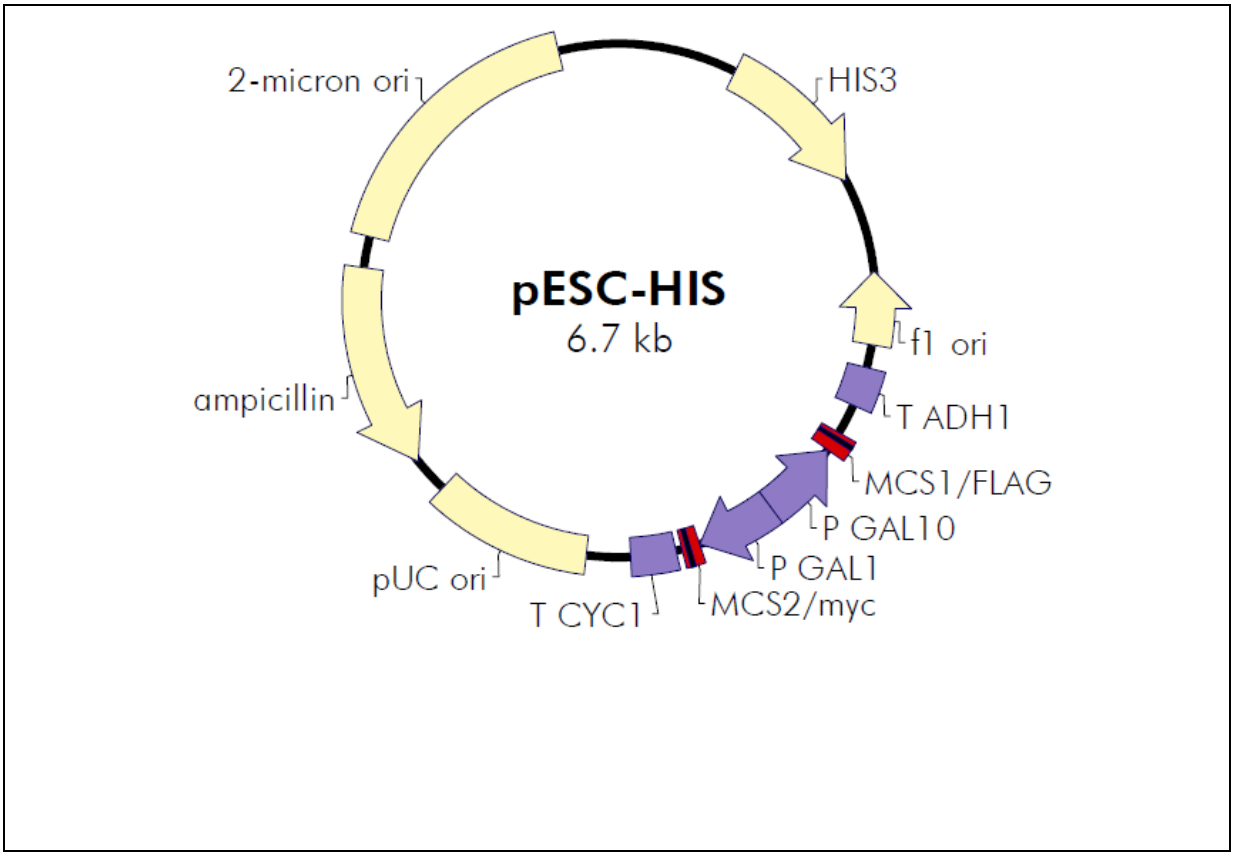



Fig. 2.3. The *HIS3* auxotrophic plasmid used in this study. This picture was obtained from the Invitrogen pESC yeast epitope tagging vectors instruction manual.

### 2.2.3. Electroporation condition

The constructed plasmids pESC-URA, pESC-TRP and pESC-HIS were separately applied as the heterologous vectors for the transformation into *P. irregulare* cells. The voltage and cuvette applied in electroporation are listed in Table 2.1. The electro-transformed *P. irregulare* cells were cultured separately in their respective trial medium described above and incubated at 26 ℃ in an orbital shaker set to 170 rpm.

Table 2.1 The voltage and cuvette applied for electroporation condition study

| | Voltage/kV with 2mm Cuvette | | | | | Voltage/kV with 4mm Cuvette | |
|---|---|---|---|---|---|---|---|
| pESC-URA | 1.0 | 1.5 | 1.8 | 2.0 | 2.5 | 1.0 | 1.5 |
| pESC-TRP | 1.0 | 1.5 | 1.8 | 2.0 | 2.5 | 1.0 | 1.5 |
| pESC-HIS | 1.0 | 1.5 | 1.8 | 2.0 | 2.5 | 1.0 | 1.5 |

### 2.2.4. Methods of analysis

#### 2.2.4.1. Biomass analysis

After harvest, *P. irregulare* cells were transferred from the culture and centrifuged at 4000 g in 50 mL sterilized tubes for 5 min. Medium was removed and the cells were washed twice with 25 mL distilled water. The pelleted fungal biomass was then freeze-dried overnight in a pre-weighed tube and re-weighed to determine biomass. After determining the weight of the dried biomass, approximately 20 mg of dried cell biomass was used for further fatty acid analyses.

#### 2.2.4.2. Fatty acid and EPA analysis

The protocol originally developed by Indarti et al (Indarti E. et al. 2005) was used for analysis of total fatty acids (TFA) and EPA. The harvested and freeze-dried *P. irregulare* cells were analyzed by gas chromatography (GC). Fatty acid methyl esters (FAME) were first extracted using a 4 mL mixture of methanol, concentrated sulfuric acid, and chloroform (1.7:0.3:2.0, v/v/v), 20 mg of dried cell biomass, and 1 mg heptadecanoic acid (C17:0) as an internal standard. No heptadecanoic acid (C17:0) existed in the TFA content of *P. irregulare* cells. In addition, with a

fixed retention time and amount, other fatty acids including EPA could be calculated and analyzed quantitatively. After incubation in water bath at 90 ℃ for 40min, 1 mL of distilled water was added into the prepared FAME solution until the mixture was cooled to room temperature. The FAME solution was then mixed completely with a vortex at least 1min and settled to allow separation of the two phases. The lower phase contained FAME and anhydrous sulfate sodium was used to absorb and remove water from the lower phase. The dried material (lipid oil phage) was transferred into an analysis vial with volume of 1.5ml. A Shimadzu 2010 gas chromatograph was used to determine the EPA and TFA content. The TFA content was obtained from the addition of all fatty acids peaks in the chromoatogrpah. The GC was equipped with a flame-ionization detector and a SGE Sol Gel-Wax$^{TM}$ capillary column (30m ×0.25mm ×0.25um). The injector was maintained at 250 ℃, and the injection amount was 1 μL by split injection mode (Ratio: 10:1). Helium was used as the carrier gas. The detector temperature was kept at 300 ℃. The column temperature profile was set as follows: 80 ℃ for 0.5 min; raised to 175 ℃ at 30 ℃ /min; raised to 260 ℃ at 5 ℃ /min; maintained for 6 min; raised to 280 ℃ at 30 ℃ /min; maintained for 1 minute. The fatty acids profile of *P. irregulare* samples was identified by comparing the retention times with those of standard fatty acids and quantified by comparing peak areas to that of the C17:0 internal standard.

## 2.3. Results

### 2.3.1. Electroporation results

The electroporation voltages studied to shock *P. irregulare* cells are given in Table 2.2. The results of transformation are also given in Table 2.2. For this research, following electroporation, the culture was grown in the liquid trial medium deficient in the respective medium component

(URA, HIS, or TRP). After transferring to solid medium, resulting colonies were then analyzed for the presence of the recombinant plasmid using PCR. The presence of plasmid under any of the electroporation conditions listed in Table 2.1 is given by "Y" (plasmid detected) and "N" (plasmid not detected). It was demonstrated that 2.0 kV was most suitable for electroporation of plasmids into the host using a 0.2 cm cuvette. It is suspected that higher voltage may have led to cell death. Other electroporation parameters were held constant in this study and included: (i) capacitance of 25 μF, (ii) resistance of 200 Ω, and (iii) pulse time of 5.0 ms.

Table. 2.2. Results of electroporation

| | Voltage/kV with 2mm Cuvette | | | | | Voltage/kV with 4mm Cuvette | |
|---|---|---|---|---|---|---|---|
| | 1.0 | 1.5 | 1.8 | 2.0 | 2.5 | 1.0 | 1.5 |
| pESC-URA | N | N | N | Y | Y | N | N |
| pESC-TRP | N | N | N | Y | N | N | N |
| pESC-HIS | N | N | N | Y | N | N | N |

"Y" (Plasmid detected), "N" (Plasmid not detected) in *P. irregulare* when analyzed by PCR.

**2.3.2. Growth characterization**

Wild-type *P. irregulare* was harvested in the form of a cloud of mycelia from the control medium consisting of glucose (30g/L) and yeast extract (10g/L). Growth curves were determined based on the dry weight of biomass, which was determined after freeze-drying. The dried biomass weight from wild-type *P. irregulare* culture reached its peak of 6.8 g/L after seven days. The growth curves were also obtained for *P. irregulare* cultures containing the auxotrophic pESC-URA, pESC-TRP, and pESC-HIS plasmids, respectively. The maximum dried cell weight

of *P. irregulare* containing pESC-URA, pESC-TRP, and pESC-HIS plasmids was 6.13 g/L, 5.3 g/L and 2.68 g/L, respectively. Growth curves are shown in Fig. 2.4.

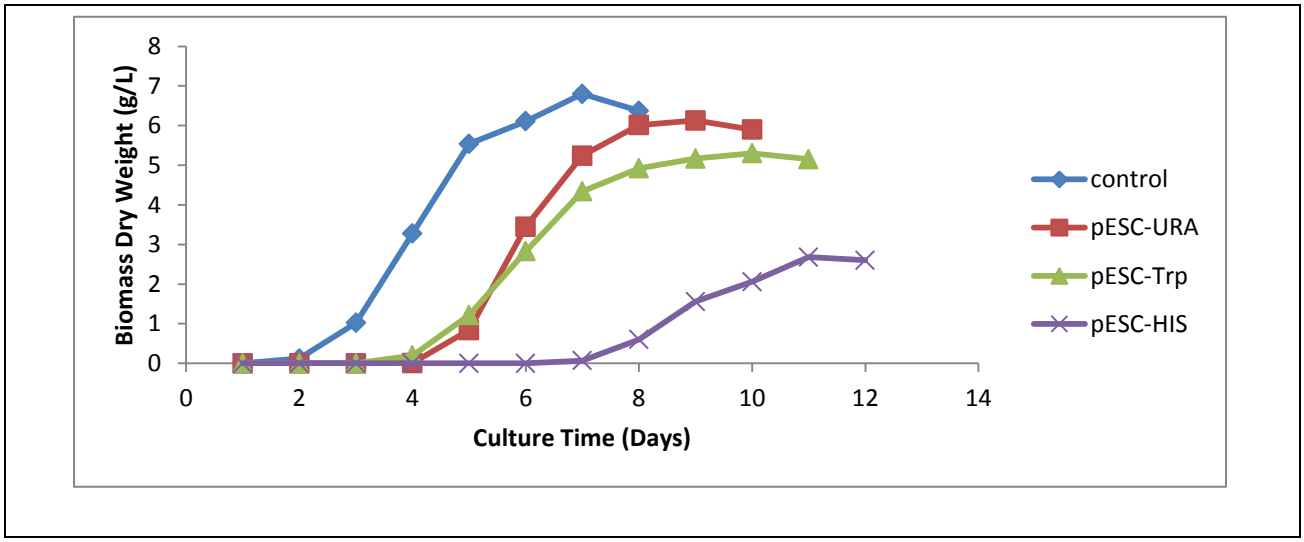



Fig. 2.4. Growth curves of *P. irregulare*: (i) wild-type (blue), (ii) containing pESC-URA (red), (iii) containing pESC-TRP (green), and (iv) containing pESC-HIS (purple).

### 2.3.3. EPA content and yield

EPA content (the weight of EPA over total dried weight of cells with the unit of mg/g) and yield (weight of EPA over the total volume of culture with the unit of mg/L) were determined for the wild-type and transformed *P. irregulare* cultures by the FAME approach using GC. All cultures were harvested at their highest level of biomass density. The maximum observed EPA content was 16.68 mg/g from *P. irregulare* containing the pESC-TRP plasmid. This increased the EPA content 52.33% relative to the EPA content of the wild-type culture. The maximum observed EPA yield was 98.52 mg/L from *P. irregulare* containing the pESC-URA plasmid. This represented a 32.28% increase relative to the wild-type. The EPA contents and yields of the wild-type and transformed *P. irregulare* cultures are given in Fig. 2.5 and table. 2.3.

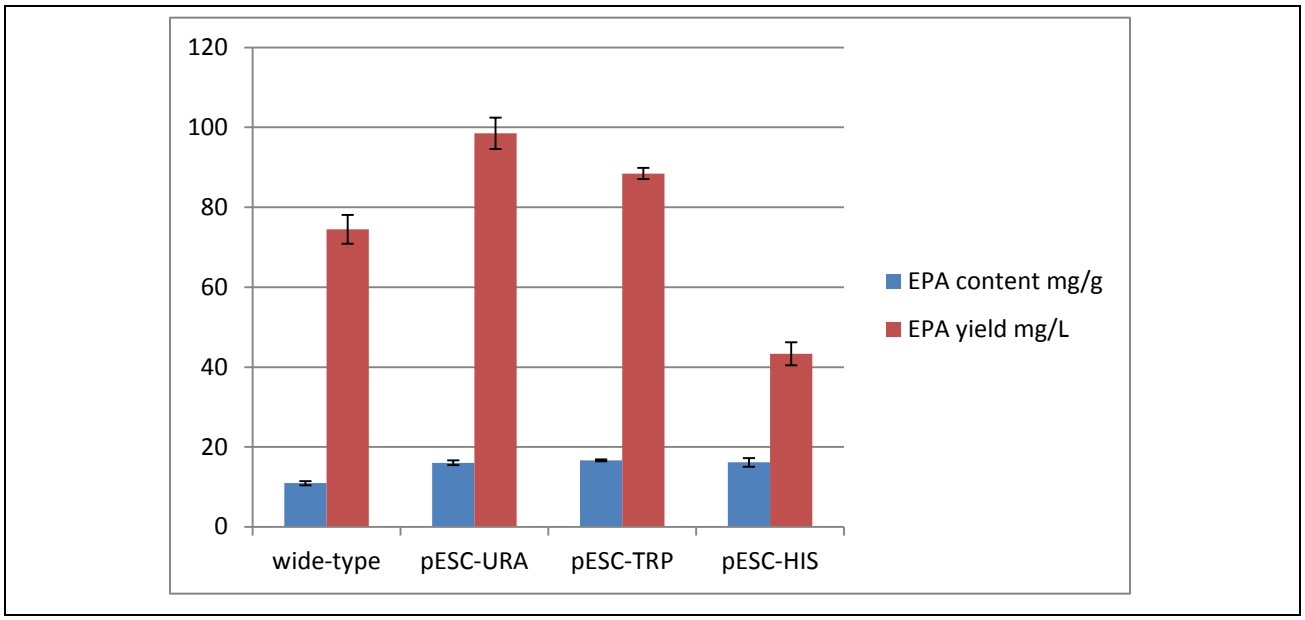



Fig.2.5. EPA content and yield for the wild-type and transformed *P. irregulare* cultures. All cultures were harvested at their highest levels of biomass concentration

**2.3.4. TFA content and yield**

The TFA content and yield were calculated from the summation of all fatty acid peaks resolved in the gas chromatographs. Fig. 2.6 and table. 2.3 illustrated the TFA content and yield of the wild-type and transformed *P. irregulare* cultures. The maximum observed TFA content was 177.26 mg/g from *P. irregulare* containing the pESC-TRP plasmid. This increased the TFA content 44.24% relative to the wild-type culture. The maximum observed TFA yield was 1026.10 mg/L from *P. irregulare* containing the pESC-URA plasmid. This represented a 22.79% increase relative to the wild-type.

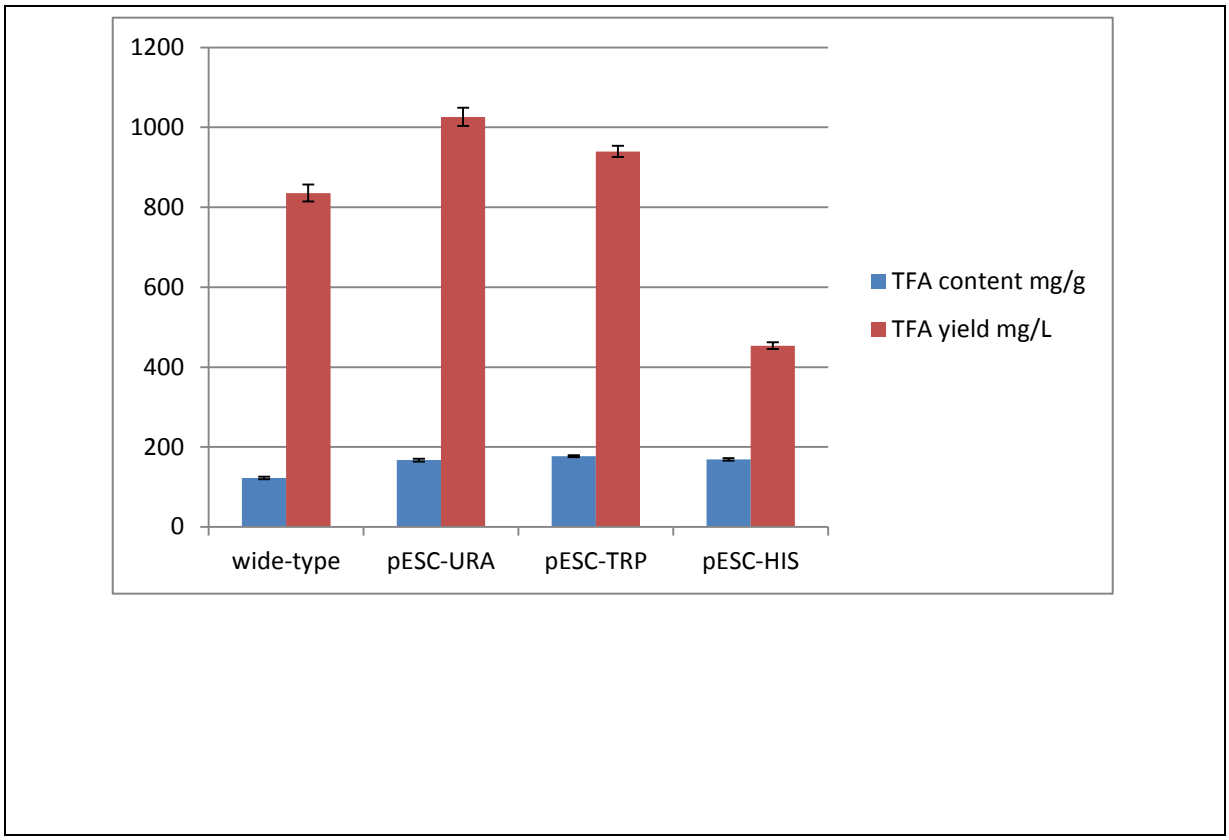



Fig. 2.6. TFA content and yield for the wild-type and transformed *P. irregulare* cultures. All cultures were harvested at their highest levels of biomass concentration.

Table. 2.3. Results of yield and content of EPA and TFA

| Strain | EPA Yield mg/L | EPA Content mg/g | TFA Yield mg/L | TFA Content mg/g |
|---|---|---|---|---|
| Wild type | 74.48 | 10.95 | 835.65 | 122.89 |
| pESC-URA | 98.52 | 16.07 | 1026.10 | 167.39 |
| pESC-HIS | 88.42 | 16.68 | 939.48 | 177.26 |
| pESC-TRP | 43.34 | 16.17 | 353.40 | 169.18 |

### 2.3.5. Content and yield of TFA

TFA included eight other fatty acids (in addition to EPA), including ALA (C20:4, n-6). Several of these eight fatty acids, such as C18:0, C18:1, and ALA, are crucial metabolites in the

biosynthesis pathway of EPA. The GC retention times for each fatty acid is shown in Fig. 2.7, along with the internal standard (C17:0). The percentage of each fatty acid in the TFA, for both content and yield, are reported in Table 2.4.

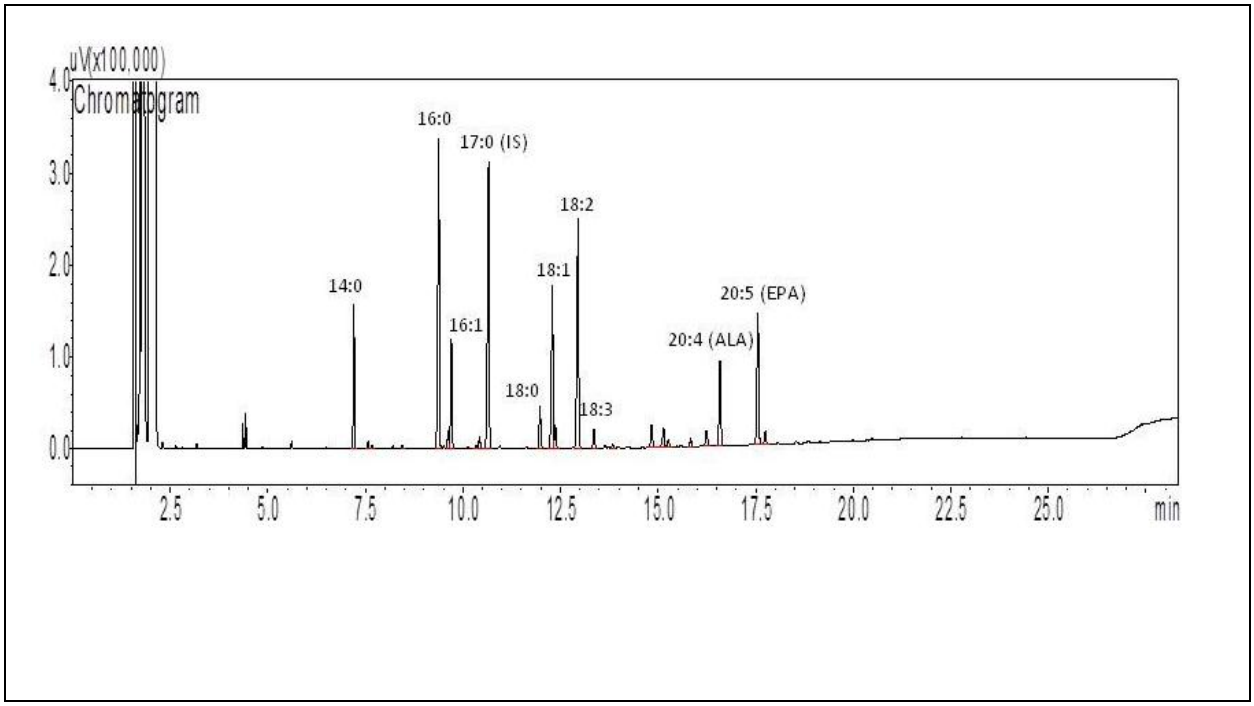



Fig.2.7. Sample gas chromatograph showing the presence of fatty acids comprising the total TFA content. The internal standard (IS) of C17:0 is also shown in this figure.

Table. 2.4. TFA content and yield

| Fatty Acid | Strain | Percent of TFA | Content (mg/g) | Yield (mg/L) |
|---|---|---|---|---|
| C14:0 | Wild-Type | 4.09 | 5.03 | 34.18 |
| | URA | 5.96 | 10.18 | 62.38 |
| | TRP | 5.38 | 9.60 | 50.90 |
| | HIS | 5.66 | 11.08 | 29.68 |
| C16:0 | Wild-Type | 17.44 | 21.43 | 145.73 |
| | URA | 17.93 | 30.63 | 187.77 |
| | TRP | 17.95 | 32.01 | 169.65 |
| | HIS | 18.32 | 35.88 | 96.15 |
| C16:1 | Wild-Type | 2.94 | 3.61 | 24.56 |

| | | | | |
|---|---|---|---|---|
| | URA | 5.59 | 9.54 | 58.51 |
| | TRP | 4.78 | 8.53 | 45.19 |
| | HIS | 5.35 | 10.47 | 28.07 |
| C18:0 | Wild-Type | 2.66 | 3.27 | 22.26 |
| | URA | 2.50 | 4.26 | 26.14 |
| | TRP | 2.67 | 4.75 | 25.20 |
| | HIS | 2.23 | 4.36 | 11.68 |
| C18:1 | Wild-Type | 9.03 | 11.10 | 75.47 |
| | URA | 10.20 | 17.42 | 106.81 |
| | TRP | 10.50 | 18.73 | 99.27 |
| | HIS | 9.79 | 19.17 | 51.36 |
| C18:2 | Wild-Type | 15.61 | 19.19 | 130.48 |
| | URA | 14.96 | 25.56 | 156.71 |
| | TRP | 14.52 | 25.89 | 137.20 |
| | HIS | 14.49 | 28.39 | 76.08 |
| C18:3 | Wild-Type | 1.42 | 1.74 | 11.84 |
| | URA | 1.12 | 1.91 | 11.73 |
| | TRP | 1.18 | 2.10 | 11.14 |
| | HIS | 1.36 | 2.67 | 7.16 |
| C20:4 | Wild-Type | 5.11 | 6.28 | 42.69 |
| | URA | 5.59 | 9.56 | 58.60 |
| | TRP | 5.21 | 9.30 | 49.27 |
| | HIS | 5.55 | 10.86 | 29.11 |

## 2.4. Discussion

### 2.4.1. Electro-transformation of P. irregulare

It was shown in this study that *P. irregulare* has the capability to accept and replicate homogeneous yeast plasmids introduced though electroporation. The growth curves, EPA, and TFA content and yield of *P. irregulare* were all impacted by the expression of the transformed plasmids. In this study, the yeast auxotrophic plasmids pESC-URA, pESC-TRP and pESC-HIS were all successfully transformed into *P. irregulare* by electroporation. The 2-micron origin of replication was shown to be compatible between yeast and *P. irregulare*. However, several false positive results were also observed along with a slowing of the growth rate of transferred *P. irregular* in trial medium. These false positive colonies survived usually survived for less than 4 days and then began to degrade themselves. This phenomenon indicated the effect of auxotrophic selection in *P. irregulare* was not as remarkable as observed in yeast, especially pESC-TRP

vector in the host. It is also possible that some competent *P. irregulare* mycelia were killed during electro-transformation and their cell lysates provided the deficient amino acids for the survival of others in auxotrophic medium. Another reason that potentially leads to a false positive result was the growth morphology. The mycelia grew in clouds and high density regions possibly blocked the medium diffusion from the outer to the inner layers. In this case, the lack of nutrients of the inner mycelia resulted in the cell death and lysis and, therefore, also offered additional amino acids for out layer mycelia. A previous study [10] indicated that decreasing rotational speed in baffled flask growth resulted in an improved pelleted morphology that ultimately benefited the accumulation of biomass. To some extent, this approach may also serve to reduce the false positive result caused by diffusion limitations. Although the electroporation efficiency was regarded as a critical factor, the mycelia morphology was suspected as causing the inaccuracy and promiscuity of single colony determination on solid media.

### 2.4.2. Auxotrophic growth characteristics

From the growth curves shown in Fig. 2.4, the maximum dried biomass of *P. irregulare* containing auxotrophic plasmids pESC-URA, pESC-TRP and pESC-HIS were 6.13 g/L, 5.3 g/L and 2.68 g/L, respectively. These were much less than that of wild-type *P. irregulare*, which was measured as 6.8 g/L. Compared to the wild-type *P. irregulare* growth curve, postponed growth was observed for electro-transformed *P. irregulare.* This was possibly due to host recovery from electroporation perforation. The reduced maximum dried biomass yield and the extended lag phases both possibly illustrate the importance of each auxotrophic amino acid in the growth of *P. irregulare.* Because the expression levels of the auxotrophic genes from the

respective plasmids were the same, histidine was regarded more essential than tryptophan and uracil to maintain the survival of *P. irregulare.* It was deduced that more L-histidine was demanded for constructing crucial proteins or that the feedback route of L-histidine biosynthesis was limited or non-existent in *P. irregulare.* In electroporation experiments, the number of successful colonies formed from pESC-HIS electroporation was much lower than what was observed for pESC-URA and pESC-TRP under identical electroporation parameters. This also indicates that L-histidine possibly plays an important role in the survival of *P. irregulare* electro-transformation. The sizes of the three plasmids were also similar.

### 2.4.3. EPA yield in auxotrophic cultures

Although it was recently reported that lower growth temperatures benefit the EPA production and accumulation [15, 16], the temperature applied in this study was kept at 26 ℃ in all experiments. The content of EPA increased dramatically with auxotrophic plasmid addition and media conditions. Increased EPA production (yield and content), of at least 45%, was observed in all strains with auxotrophic plasmids compared to wild-type *P. irregulare*. The observed enhancement of EPA production indicated altered metabolism and cellular phenotype elicited by auxotrophic growth with URA, TRP and HIS, respectively. However, it was intriguing that the shortage of these three amino acids resulted in an accumulation of not only EPA but also other polyunsaturated fatty acids (PUFA). In general, the biosynthesis of PUFA including EPA in fungus involves aerobic pathways to convert LA (18:2 Δ9,12) or ALA (18:3, Δ9,12,15) into EPA (20:5 Δ5,8,11,14,17) though either a ω-3 or ω-6 route [17]. Because of the high activity of Δ12 desaturase that converts oleic acid into linoleic acid and Δ17 desaturase that converts arachidonic acid (AA) into EPA, *P. irregulare* is considered to synthesize EPA mainly through

the ω-6 route. The evidence that an increasing percentage of each PUFA distribution were respectively detected by GC from stearic acid (C18:0) to AA (C20:4) (Table 2.2) also supported this hypothesis. The activity of the Δ17 desaturase also catalyzes AA into EPA via the ω-3 route. The decreased percentage of GLA (C18:3) in the three *P. irregulare* strains with auxotrophic plasmids indicates a possible increase in activity of the Δ6 elongase in the auxotrophic media.

### 2.4.4. Possible mechanisms leading to altered EPA yield and content

Both the content and the yield of EPA and other PUFA were enhanced in the *P. irregulare* strains with auxotrophic plasmids cultured in appropriate auxotroph medium. This phenomenon implied a close relationship between PUFA biosynthesis and the biosynthetic metabolic pathways of these amino acids. Although this exploratory research will lead the way for additional findings, it is suspected that the metabolomic change in the amount of free acetyl-CoA *in vivo* caused by enriching deficient essential amino acids is tied directly to changes in PUFA biosynthesis. The desaturation and elongation occurred with the acyl cycling between phosphatidylcholine pool and the CoA pool [18]. An increase in the availability of acyl possibly resulted in an enhancement of the content of stearic acid (C18:0) and then further improved the content of EPA. However, the dramatic increase of the percentage of fatty acid C16:1 in *P. irregulare* strains with auxotrophic plasmids might imply yet another biosynthetic route to EPA.

### 2.4.5. Conclusions

In summary, the electroporation conditions and the content and yield of EPA and TFA in *P. irregulare* were determined in this research. This study has proven the feasibility of electro-

transforming and replicating heterologous yeast vectors in *P. irregulare*. All results presented in this study also indicate a great potential exists to increase the EPA yield by the fermentation of *P. irregulare* strains under auxotrophic conditions. However, the (i) optimization of fermentation conditions, such as lower temperature, (ii) incubation and pelleting of mycelia, as well as (iii) three governing genes known to impact EPA yield (Δ5 desaturase, Δ6 desaturase and Δ6 elongases) were held constant in this research  This further suggests that significant improvements to EPA content and yield are possible through metabolic engineering and bioprocess engineering steps in future research.

**2.5. References**

[1] Johnson, D. T.; Taconi, K. A. The glycerin glut: options for the value-added conversion of crude glycerol resulting from biodiesel production. EnViron. Prog. 2007, 26, 338–348.

[2] Thompson, J. C.; He, B. B. Characterization of crude glycerol from biodiesel production from multiple feedstocks. Appl. Eng. Agric. 2006, 22, 261–265.

[3] Simopoulos, A. P. Essential fatty acids in health and chronic disease. Am. J. Clin. Nutr. 1999, 70, 560S–569S.

[4]Ursin, V. M. Modification of plant lipids for human health: development of functional land-based omega-3 fatty acids. J. Nutr. 2003, 133, 4271–4274.

[5] Innis, S. M. Fatty acids and early human development. Early Hum. DeV. 2007, 83, 761–766

[6] Hooper L., R. L. Thompson, R. A. Harrison, C. D. Summerbell, A. R. Ness, H. J. Moore, H. V. Worthington, P. N. Durrington, J. P. T. Higgins, N. E. Capps, R. A. Riemersma, S. B. J.

Ebrahim, and G. D. Smith. 2006. Risks and benefits of omega-3 fats for mortality, cardiovascular disease, and cancer: systematic review. British Medical Journal. 332(7544):752-760.

[7] Pyle, D. J., R. A. Garcia, and Z. Wen. 2008. Producing Docosahexaenoic Acid (DHA)-Rich Algae from Biodiesel-Derived Crude Glycerol: Effects of Impurities on DHA Production and Algal Biomass Composition. Journal of Agricultural and Food Chemistry. 56: 3933-3939.

[8] Covington, M. B. 2004. Omega-3 Fatty Acids. American Family Physician. 70(1): 133-140.

[9] Parker G, Gibson N.G., Brotchie H., Heruc G., Rees A., Hadzi-Pavlovic D. (2006) Omega-3 Fatty Acids and Mood Disorders. American Journal of Psychiatry, 163, 969-978

[10] SNEHA K. ATHALYE, RAFAEL A. GARCIA, AND ZHIYOU WEN. Use of Biodiesel-Derived Crude Glycerol for Producing Eicosapentaenoic Acid (EPA) by the Fungus Pythium irregulare. J. Agric. Food Chem. 2009, 57, 2739–2744.

[11] Rymowicz W, Rywinskaa A. and Zarowska B. (2007) Biosynthesis of citric acid from crude glycerol by Yarrowia lipolytica in repeated-batch cultivations. Journal of Biotechnology, 131, 149-150.

[12] GBIF. http://data.gbif.org/species/Global; Biodiversity Information Facility, 2008.

[13]Worthington, P. N. Durrington, J. P. T. Higgins, N. E. Capps, R. A. Riemersma, S. B. J.Ebrahim, and G. D. Smith. 2006. Risks and benefits of omega-3 fats for mortality, cardiovascular disease, and cancer: systematic review. British Medical Journal. 332(7544): 752-760.

[14] O'Brien, D. J., M. J. Kurantz, and R. Kwoczak. 1993. Production of eicosapentaenoic acid by the filamentous fungus Pythium irregulare. Applied Microbiology and Biotechnology. 40: 211-214

[15] Stinson, E. E., R. Kwoczak, and M. J. Kurantz. 1991. Effect of cultural conditions on production of eicosapentaenoic acid by Pythium irregulare. Journal of Industrial Microbiology. 8(3): 171-178.

[16] Liao, W., Y. Liu, C. Frear, and S. Chen. 2007. A new approach of pellet formation of a filamentous fungus – Rhizopus oryzae. Bioresource Technology. 98(18): 3415–3423.

[17]Indarti E. MM, R. Hashim, and A. Chong. : (2005) Direct FAME synthesis for rapid total lipid analysis from fish oil and cod liver oil. Journal of Food Composition and Analysis, 18, 161-170.

[18] Surinder P Singh, Xue-Rong Zhou, et. al. Metabolic engineering of new fatty acids in plants. Current Opinion in Plant Biology 2005, 8:197–203

## 3. Designing fine-tuned antisense RNA constructs for prokaryotes using a combinatorial approach and thermodynamics

### 3.1. Introduction and background

#### 3.1.1. Antisense technology

As a tool for gene function analysis and therapeutic application, antisense technology was first described in 1987 [1]. From there, it developed rapidly in eukaryotes, but it has remained relatively under-developed in prokaryotes due to the relative ease of gene knockouts through recombination methods. Ultimately, asRNA is of interest as a metabolic engineering tool to down-regulate enzymes participating in undesired pathways. In eukaryotic systems, the commercial application of antisense technology has enabled clinical developments as well as cellular engineering. Indeed, one antisense-based biopharmaceutical is on market for the treatment of cytomegalovirus retinitis and several are under clinical trials [2]. The utilization of antisense technology to control cell phenotype in tissue engineering is also now well-developed and is being studied to for its ability to modify biosynthetic pathways and alter primary and secondary metabolic routes for engineering cell growth and protein synthesis capacity [3]. In these applications, (i) antisense oligonucleotides, (ii) antisense RNA (asRNA), and (iii) small interfering RNA (siRNA) are three main methods to apply antisense technology in hosts. Generally, the mechanism of the antisense approach is to silence or degrade expression of a specific gene through the hybridization of a designed RNA or single-stranded oligonucleotide with a targeted mRNA transcript. The affinity of these molecules is governed by thermodynamics and the nucleic acids complementary rule [21]. However, the experimental

methods and results differ among these three antisense technologies. Fig.3.1 [3] indicates the various means to produce the specific inhibition of gene expression. Considerable progress in antisense technology has been made in the past two decades. However, crucial challenges still exist in achieving "fine-tuned" gene expression through application of antisense technologies. As for antisense oligonucleotides, the inefficient delivery to cells and unfavorable intracellular trafficking leads to poor effects on suppressing gene expression when applied in cell culture [4]. Although RNA interference (RNAi) has proven to be a powerful technique on genetic regulatory circuits, it is mainly applied to eukaryotes. The use of asRNA has proven useful in prokaryotes, particularly to clostridia before gene knockout technologies were developed [5]. However, certain asRNA constructs have been found in the literature to be more effective at down-regulation than others. To this point, a guideline for asRNA design had been formulated [5], but a thermodynamic model is sought to "fine-tune" asRNA effectiveness. This has direct impacts for modern metabolic engineering strategies that seek to knockdown gene expression or pathway usage by a defined amount, without resorting to a complete knockout. This may also result in the development of metabolic switches, where pathway usage can be knocked down by the "on/off" mechanism of an inducible promoter controlling an asRNA open reading frame.

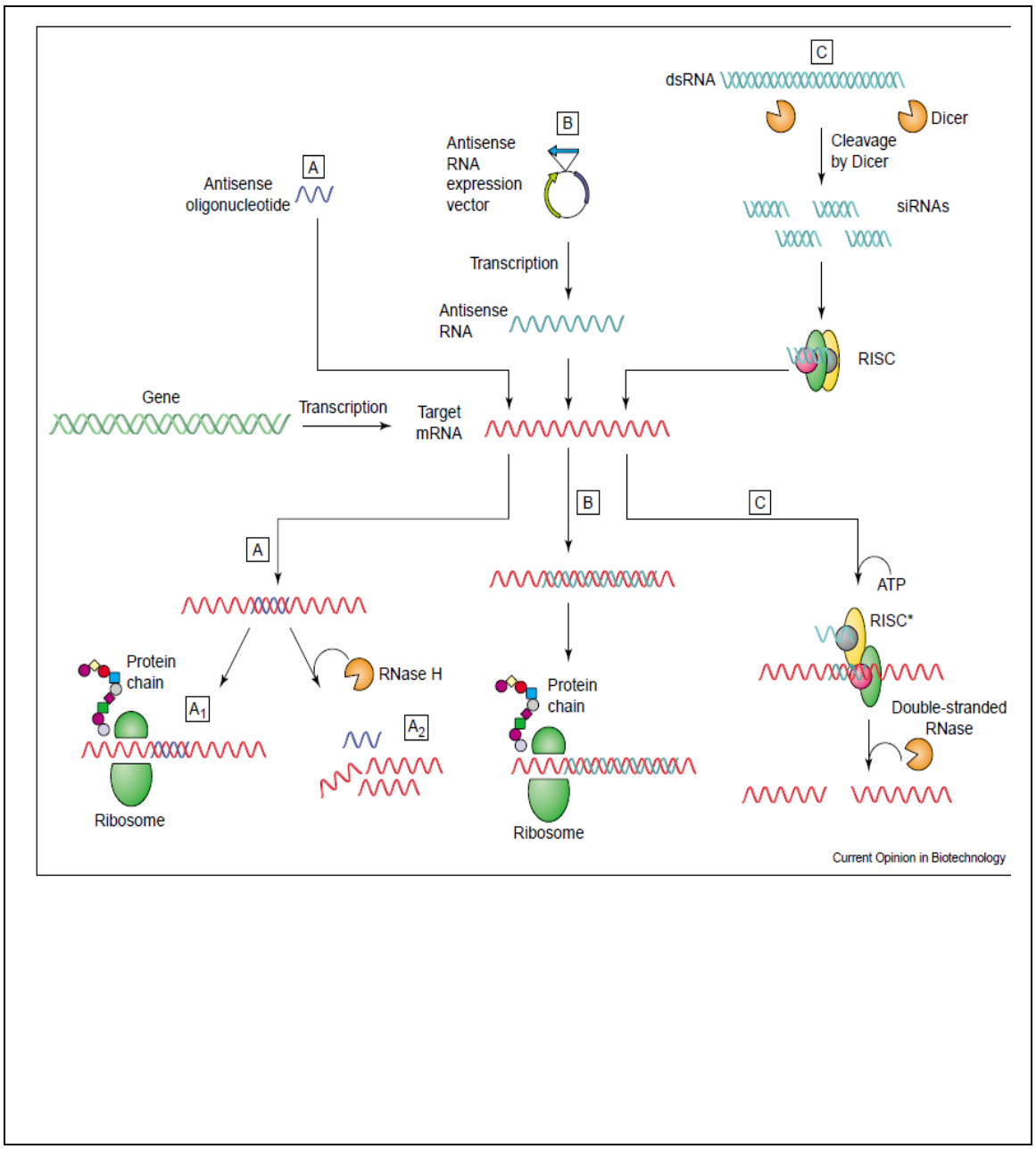



Fig. 3.1. Mechanisms of antisense technologies on the inhibition of gene expression. Track A: antisense oligonucleotide process; Track B: antisense RNA process; Track C: in RNA interference process [3].

### 3.1.2. Antisense RNA

The asRNA technology is now identified as an effective and specific method for regulating microbial gene expression at the posttranscriptional level [2]. The approach is to introduce a

recombinant expression vector into a microbial host cell with a recombinant asRNA open reading frame under control of a strong constitutive promoter. Once the asRNA is transcribed, it will specifically inhibit the synthesis of the target protein by binding the mRNA from which it is derived, and then inhibiting translation. During the asRNA-mRNA binding process, it is hypothesized that a duplex forms is formed between the asRNA and the complementary mRNA molecule. This prevents the progress of the translation of mRNA, as the ribosomal RNA is blocked during translation. It is also hypothesized that the asRNA-mRNA duplex is formed between loop structures on separate RNA molecules initiated by hydrogen bonding. This process is believed to be relatively slow and the rate-limiting step of asRNA-mRNA hybridization. The unfolding of these RNA structures is required first for hybridization, but it is hypothesized this is initiated at complimentary loop structures of the asRNA and mRNA secondary structures. Based on this, a minimum free energy exists for hybridization to occur.

Several approaches exist to transport vectors containing asRNA open reading frames into hosts: (i) electroporation [5], (ii) positively charged carrier molecules, and (iii) viral vectors such as retrovirus [6] and adenovirus [7]. A great advantage of the asRNA strategy is the capacity to design vectors with inducible transcriptional activation [8]. With an inducible promoter, this provides an engineer the capability of turning the asRNA regulation "on/off" at will. Once the promoter is active, asRNA consequently regulates the expression level of target gene (mRNA). In a metabolic engineering strategy, the substrate to the reaction catalyzed by the target enzyme (encoded by the target mRNA) may be re-routed (partially or fully) into an engineered pathway to yield a desired products. Thus, the asRNA provides the capability to turn "on/off" a metabolic engineering strategy at will, enabling physical control over metabolism of engineered bacteria. This is an additional inducible metabolic engineering tool to gene over-expression. Other tools

such as gene knockouts are neither inducible nor tunable (it is not possible to partially knockout a gene). The fact that expression of asRNA of *ctfB* gene in *Clostridium acetobutylicum* improves the fermentation of various sugars into acetone and butanol [9] also shows compelling evidence on the value of antisense RNA application.

### 3.1.3. Examples of asRNA in prokaryotes

Many asRNA-regulated gene systems in prokaryotes have been studied in detail [21]. The asRNA interaction with a target mRNA in bacteria generally involves linear-linear, loop-linear and loop-loop pairing schemes [22]. In addition,two classes of prokaryotic asRNAs can be distinguished: (i) cis-encoded and (ii) trans-encoded asRNA. The cis-encoded RNAs are located in the same DNA region and are fully complementary to their targets while trans-encoded RNAs are located in another chromosomal location and are only partially complementary to their target mRNA(s). The majority of cis-encoded antisense RNAs have been found in plasmids, phages and transposons [23] and studied as the main tool for gene regulation on expression level. The antisense *ctfB* RNA application was successful in increasing the butanol/acetone ratio in *Clostridium acetobutylicum* fermentation. The antisense RNA sequence was commonly designed to be complementary to the sequence of crucial gene and located on a heterologous vector. Its function was implemented after transcription.

### 3.1.4. Fluorescent proteins as reporters

With the first identification of green fluorescent protein (GFP) from the *Aequorea victoria* jellyfish [10] in the early 1960's, the modern era of biochemical reporting was born. Several useful approaches fuse the GFP to a wide variety of protein and enzyme targets by genetic

manipulation. This approach is rapidly expanding and is often used to track the location of a target protein or enzyme. A vast number of fluorescent proteins from other organisms and their differently-colored mutants and fusion proteins have been identified and isolated, resulting in a dramatic expansion of the color palette in recent years [10]. Various fluorescent proteins in this list promise a wide variety of options for developing new tools for gene reporting and metabolic engineering. The comprehensive list and properties of the commonly used fluorescent proteins are given in Table 3.1.

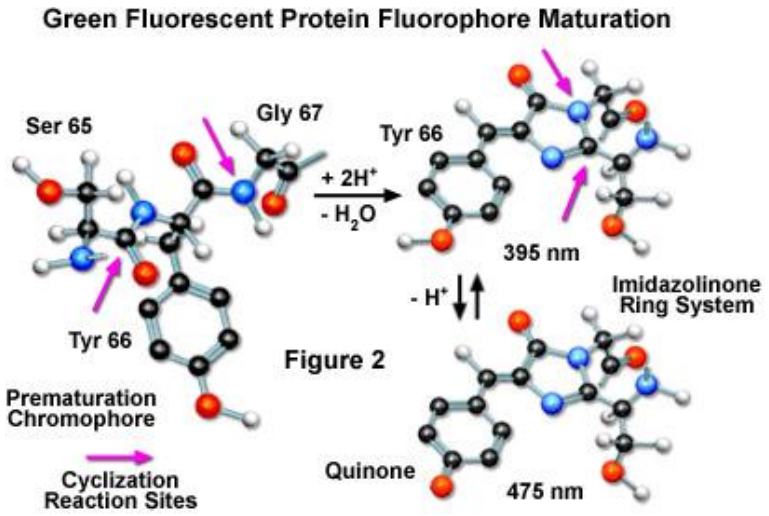
Table 3.1. Fluorescent proteins and ther properties.
Cited form http://www.microscopyu.com/articles/livecellimaging/fpintro.html

| Protein (Acronym) | Excitation Maximum (nm) | Emission Maximum (nm) | Molar Extinction Coefficient | Quantum Yield | in vivo Structure | Relative Brightness (% of EGFP) |
|---|---|---|---|---|---|---|
| GFP (wt) | 395/475 | 509 | 21,000 | 0.77 | Monomer | 48 |
| Green Fluorescent Proteins | | | | | | |
| EGFP | 484 | 507 | 56,000 | 0.60 | Monomer | 100 |
| Emerald | 487 | 509 | 57,500 | 0.68 | Monomer | 116 |
| Superfolder GFP | 485 | 510 | 83,300 | 0.65 | Monomer | 160 |
| Azami Green | 492 | 505 | 55,000 | 0.74 | Monomer | 121 |
| Blue Fluorescent Proteins | | | | | | |
| EBFP | 383 | 445 | 29,000 | 0.31 | Monomer | 27 |
| EBFP2 | 383 | 448 | 32,000 | 0.56 | Monomer | 53 |
| Azurite | 384 | 450 | 26,200 | 0.55 | Monomer | 43 |
| mTagBFP | 399 | 456 | 52,000 | 0.63 | Monomer | 98 |
| Cyan Fluorescent Proteins | | | | | | |
| ECFP | 439 | 476 | 32,500 | 0.40 | Monomer | 39 |
| mECFP | 433 | 475 | 32,500 | 0.40 | Monomer | 39 |
| Cerulean | 433 | 475 | 43,000 | 0.62 | Monomer | 79 |
| CyPet | 435 | 477 | 35,000 | 0.51 | Monomer | 53 |
| Yellow Fluorescent Proteins | | | | | | |
| EYFP | 514 | 527 | 83,400 | 0.61 | Monomer | 151 |
| Topaz | 514 | 527 | 94,500 | 0.60 | Monomer | 169 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Venus | 515 | 528 | 92,200 | 0.57 | Monomer | 156 |
| mCitrine | 516 | 529 | 77,000 | 0.76 | Monomer | 174 |
| Orange Fluorescent Proteins | | | | | | |
| Kusabira Orange | 548 | 559 | 51,600 | 0.60 | Monomer | 92 |
| Kusabira Orange2 | 551 | 565 | 63,800 | 0.62 | Monomer | 118 |
| mOrange | 548 | 562 | 71,000 | 0.69 | Monomer | 146 |
| mOrange2 | 549 | 565 | 58,000 | 0.60 | Monomer | 104 |
| Red Fluorescent Proteins | | | | | | |
| mRuby | 558 | 605 | 112,000 | 0.35 | Monomer | 117 |
| mApple | 568 | 592 | 75,000 | 0.49 | Monomer | 109 |
| mStrawberry | 574 | 596 | 90,000 | 0.29 | Monomer | 78 |
| AsRed2 | 576 | 592 | 56,200 | 0.05 | Tetramer | 8 |
| mCherry | 587 | 610 | 72,000 | 0.22 | Monomer | 47 |

### 3.1.5. Specifics about the GFP

The GFP is composed of 238 amino acids and emits bright green fluorescence when exposed to blue light (ex = 480 nm, em = 515 nm; Table 3.1) The fluorophore originates from an internal Ser-Tyr-Gly sequence, which is post-translationally modified to a 4-(p-hydroxybenzylidene)-imidazolidin-5- one structure [11]. The GFP tertiary structure and its maturation process are shown in Fig. 3.2.




Cited from http://www.microscopyu.com/articles/livecellimaging/fpintro.html

Fig. 3.2. The GFP fluorophore maturation process.

### 3.1.6. mCherry fluorescent protein

The mCherry fluorescent protein (ex = 587 nm, em = 610 nm; Table 3.1) is a red fluorescent protein and a variant from DsRed, the first coral-derived fluorescent protein to be utilized extensively. The mCherry was derived because of the disadvantage with DsRed that fluorescence occurs slowly and proceeds through a time period when fluorescence emission is in the green region. Variants were explored to enhance the effective reporting capabilities of this protein in the red emission region. The mCherry mutation overcomes the shortcomings of instability of DsRed and also leads the excitation and emission peaks red-shift in both visible light and fluorescence [12, 13].

### 3.1.7. Thermodynamics and the Nucleic Acid Package (NUPACK) Software

After the accumulation of a vast dataset of RNA sequence and structural information, sequence-structure relationships have been developed for RNA mechanisms of catalysis, regulation, and information storage [14, 15]. The implementation of computer programs and algorithms for RNA secondary structure predictions has been developed in last several years [12]. These are based upon the principle that RNA secondary structure is determined from its primary sequence by Gibbs free energy minimization [21]. Three trends have emerged from this: (i) applying statistical mechanics of RNA folding [22], (ii) using algorithms that allow pseudoknots [16], and (iii) finding the secondary structure common to a set of homologous sequences [15]. Based on these methods, a large number of predicted secondary structures with the lowest Gibbs free energy structure were analyzed, with a given sequence. Essentially, these possible structures are considered to be implicit on dynamic programming algorithms [16,17] without experimental support. Only when combined with the various known stable motif/loop structures of RNA, as well as new parameters to the model internal loops, is the prediction of RNA secondary structures feasible under dilute solution conditions. A summary of recently developed RNA secondary structure prediction software [18] is provided as Table 3.2.

Table 3.2. Summary of recently developed RNA secondary structure prediction algorithms and software

| Summary of recently developed RNA secondary structure prediction software | | |
|---|---|---|
| Advance | Program | URL |
| -Using color annotation of base pair probabilities to determine base pairs predicted with | RNA structure | http://ma.urmc.rochester.edu |

| | | |
|---|---|---|
| high confidence | | |
| -Secondary structure prediction by sampling | S Fold | http://sfold.wadsworth.org<br>http://www.bioinfo.rpi.edu/applications/sfold |
| -Iterated loop matching prediction of pseudoknots | ILM | http://www.cse.wustl.edu/~zhang/projects/ma/ilm |
| -Heuristic prediction of pseudoknots | Hot Knots | http://www.cs.ubc.ca/labs/beta/Software/HotKnots |
| -Determination of a secondary structure common to multiple sequences using abstract shapes | RNA cast | http://bibiserv.techfak.uni-bielefeld.de/rnacast/supplementary.html |
| -Prediction of RNA secondary structure common to two unaligned sequences | FOLDALIGN | http://foldalign.kvl.dk |
| -Prediction of RNA secondary structure common to two unaligned sequences | Dynalign | http://ma.urmc.rochester.edu |
| -Discovery of ncRNA genes | RNAz | http://www.tbi.univie.ac.at/~wash/RNAz |
| This table summarizes the major developments in secondary structure prediction, lists the implementing program, and provides a URL for downloading or using the software online. | | |

The freely available nucleic acid package (NUPACK) (Microfost), which is also offered from a web server (http://www.nupack.org), is intended for the analysis and design of nucleic acid secondary structure systems and contains three powerful functions: (i) analysis, (ii) design and (iii) utilities [19]. NUPACK algorithms are formulated in terms of ordered complexes. In most cases, pseudoknots are excluded from the structural ensemble. The interaction of nucleic acids involving one or more species of strands formed un-pseudoknotted complexes in dilute solution could be analyzed based on the input reaction conditions such as: (i) temperature (or range of temperatures for melts), (ii) number of strand species, (iii) maximum complex size, (iv) strand

sequences, and (v) strand concentrations. The partition function, equilibrium base-pairing probabilities, and minimum free energy (MFE) structure are calculated using dynamic programs [19, 20]. This dynamic program is based on a rigorous extension of secondary structure models to the multistranded case, addressing representation and distinguishability issues that do not arise for single-stranded structures [24]. According to the program, the secondary structure of hybridization complex, minimum free energy and combination percentage of antisense RNA were able to be calculated. The algorithm of the thermodynamic analysis of interacting nucleic acid strands was based on the formation of base pairs between complementary bases (Watson–

Crick pairs C-G and A-T (DNA) or A-U (RNA) and less frequent pairs G-T (DNA) or G-U (RNA)) [24]. For a given sequence, the free energy of secondary structures is estimated as the sum of the empirically determined free energies of the constituent loops [25, 26, 27]. The secondary structure was based on the most stable complex. The design function of the NUPACK software allows the design of nucleic acid sequences for complexes, which are intended to adopt a target secondary structure at equilibrium. The equilibrium properties of a complex could also be evaluated and displayed by the function of utilities.

### 3.1.8. Purpose of this research

This research intended to use thermodynamics to design effective asRNA constructs. A combinatorial PCR approach was designed to produce several asRNA fragments to the fluorescent *GFP* and *mCherry* mRNA. This research also investigated *lacZ* as well. Different asRNA constructs were found to knockdown gene expression of these reporters differently. Thermodynamic calculations were then made for each asRNA and target mRNA hybridization and relationships were sought with the observed expression level. This is the first model to allow the calculation of a reduced gene expression level based on thermodynamic calculations.

## 3.2. Materials and Methods

### 3.2.1. Strains, plasmids and reagents

The *Escherichia coli* Top10 strain ( *F- mcrA Δ*(*mrr-hsdRMS-mcrBC*) *φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ*(*ara-leu*) *7697 galU galK rpsL* (*StrR*) *endA1 nupG λ*-) (Invitrogen) was used as the host for cloning and replicating recombinant plasmids. The *E. coli* 10-beta (*araD139* Δ(*ara-leu*)7697 fhuA lacX74 galK (*ϕ80* Δ(*lacZ*)*M15*) *mcrA galU recA1 endA1 nupG rpsL* (*StrR*) Δ(*mrr-hsdRMS-mcrBC*))strain (New England Biolabs) and *E.coli* BL21 (DE3) (fhuA2 *[*lon*]* ompT gal *(*λ DE3*) [*dcm*]* ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::*(*lacI::PlacUV5::T7 gene1*)* i21 Δnin5*)* were used for expressing heterologous genes. Liquid Luria Bertani (LB) medium and LB agar plates were prepared according to the following protocol to culture *E. coli*: (i) 10 g tryptone, 5g yeast extract and 10g NaCl (15g agar if solid plates) were dissolved in 800 mL deionized water, (ii) pH adjusted to 7.0 and total volume to 1 L with deionized water, (iii) autoclaved at 121 ℃ for 25 min, (iv) added antibiotics or chemical indictors (X-gal and IPTG) in proper concentrations after the medium was cooled to about 55 ℃, (v) for solid media, 20 mL of liquid medium was placed into each plate, the medium solidified at room temperature, and the plate was sealed with parafilm, and (vi) the liquid media and plates were stored +4 ℃. Both strains (Top10 and 10-bata) of *E. coli* were grown in liquid LB medium at 37 ℃ overnight at 220 rpm. A finial concentration of 100 μg/mL of ampicillin was added if necessary. The common pUC19 plasmid (Invitrogen) was used as the backbone for all plasmid construction and gene expression. The *VentR* proof-reading polymerase (New England Biolabs) was used for gene amplification by PCR. All PCR reagents, antarctic phosphatase and other digested enzymes (*HindIII, PstI, BglII,Eco0109I, AatII, xhoI and NdeI*) were also obtained from New England Biolabs. All PCR primers were synthesized by Integrated DNA Technologies (IDT). PCR primers were designed using Primer3Plus software (http://www.bioinformatics.nl/cgi-

bin/primer3plus/primer3plus.cgi). All PCR primers designed and used in this research are given in Table 3.3. Beta-galactosidase enzyme assay system kit was purchased from Promega. The thermodynamic analysis of asRNA and mRNA complexes was implemented by NUPACK software offered from http://www.nupack.org.

### 3.2.2. Genetic manipulations

The extraction of genomic DNA and plasmids from *E. coli* was carried out after harvest, using a Generation Capture Kit (Promega) according to the manufacturer's protocol. PCR products were purified using a PCR purification kit (QIAGEN) (Cat. No 28104) according to manufacturer's protocol. In selected cases, gel-band purification was performed using SYBR-safe dye (Invitrogen) and a kit (QIAGEN) (Cat. No 28704) according to the manufacturer's protocol. All experimental processes were according to standard protocol unless additional stated.

#### 3.2.2.1. Promoter and terminator amplification from pUC19

The *plac* promoter and rho independent terminator of the *lacZ* gene of the pUC19 plasmid were cloned respectively using their designed primer sets (Table 3.3) and the following PCR reaction components with the pUC19 plasmid as the DNA template: (i) 5 µL 10x thermo buffer, (ii) 2.5 µL forward and reverse primers (1.0 ng/ml) ), (iii) 1 µL template DNA (plasmid at (10pg-10ng/ml))), (iv) 1 µL stock dNTP mixture, (v) 1 µL VentR polymerase, (vi) 37 µL molecular biology grade water (total volume was 50 µL). The primers used for the *plac* promoter were a touch-down PCR program was applied and is described as follows: (i) denaturing at 94 ℃ for 5 min, (ii) annealing at a specified temperature for 1 min, and (iii) elongation at 72°C for 1.5 min. The annealing temperatures used in the touch-down procedure were (i) 56°C (5 cycles), (ii) 54°C

(5 cycles), (iii) 52°C (5 cycles), and (iv) 50°C (15 cycles). A final extension step at 72 ℃ for 7 min was also applied. The PCR products were visualized by 2% agarose gel and were recovered by gel band purification for cloning (QIAGEN).

### 3.2.2.2. Construction of the pUC19-*pro-term* plasmid

The pUC19 plasmid was extracted from harvested *E. coli* and digested with *Eco0109I* and *AatII* restriction enzymes. The digestion reaction contained the following: (i) 20 µL pUC19 plasmid (10pg-10ng/ml), (ii) 6 µL 10XBuffer 4, (iii) 6 µL 10XBSA, (iv) 1 µL *Eco0109I*, (v) 1 µL *AatII*, and (vi) 26 µL deionized water (total volue was 60 µL). The digestion was incubated at 37 ℃ overnight. Then the digested plasmid was visualized and recovered by gel-purification and dissolved in with 50 µL molecular biology grade water. The recovered linearized plasmid was then ready for ligation. The *plac* promoter and *lacZ* terminator PCR products were also digested, separately, with *Eco0109I/ BglII* and *BglII/AatII*, respectively. After purification, these two DNA fragments were mixed at the ratio of 1:1 and ligated to each other with T4 ligase at 16 ℃ overnight. The ligased DNA product was gel-band purified, and this ligation product was further ligated into the linearized pUC19 plasmid, which had digested with *Eco0109I* and *AatII*. A molar ratio of 3:1 (insert:plasmid) was used for the ligation, and incubation occurred at 16 ℃ overnight. The resulting plasmid is shown in Fig. 3.3.

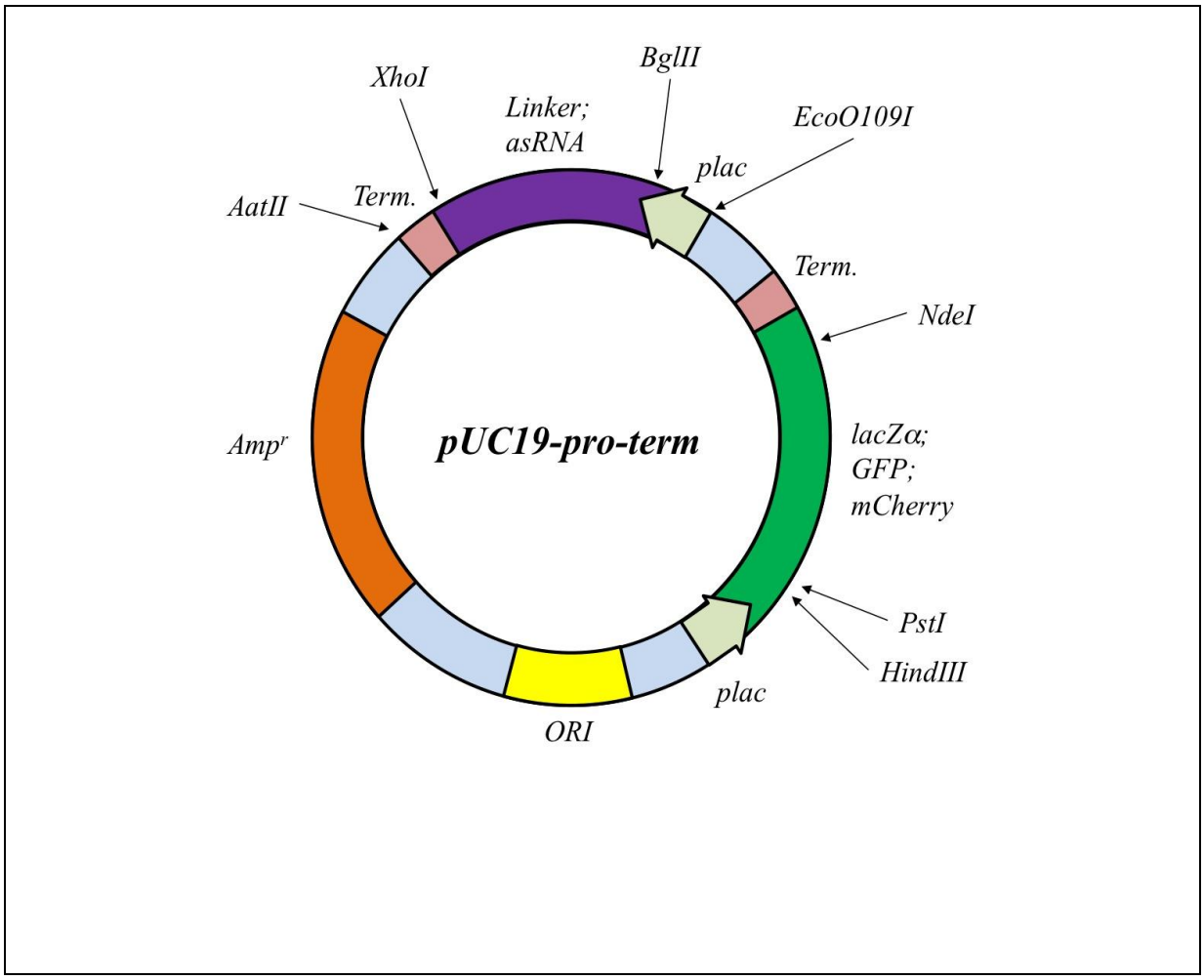



Fig. 3.3. Diagram of the pUC19-*pro-term* plasmid with enzyme restriction sites and additional inserts.

### 3.2.2.3. Chemical transformation into *E. coli Top10* cells

To perform the chemical transformation, 10 µL of ligation solution was added into one 1.5 mL vial of chemically competent *E. coli Top10* cells. After a 30 min incubation on ice, the mixed culture was heat shocked at 42 ℃ for 30 seconds. Then, the culture was incubated on ice again for 2 min. Then, 1 mL of SOC medium (20 g tryptone, 5 g yeast extract, 2mL of 5 M NaCl, 2.5 mL of 1 M KCl, 10mL of 1 M MgCl2, 10mL of 1M MgSO4, 20ml of 1 M glucose) was added to the mixture and cultured at 37 ℃ for 1 hour. Then, 100 µL of the culture were plated on solid LB media containing 100 µg/mL ampicillin and 100 µL X-gal/IPTC. After incubation overnight at

37 ℃, white colonies were extracted and sub-cultured in 5 mL liquid LB medium through exponential growth. The culture was harvested, and the plasmid was extracted as previously described.

### 3.2.2.4. Construction of *puc19-pro-term-aslacZ*

The *lacZ* gene was used as the template for designing asRNA to the *lacZ* gene. When cloned into the plasmid, these are referred to as asRNA genes. Primers are listed in Table 3.3. A total of 19 asRNA genes were PCR amplified and purified with the PCR purification kit. The PCR fragments were then digested with *XhoI* and *BglII* restriction enzymes, with incubation overnight at 37 ℃. The digested PCR fragments were separately mixed in a 3:1 (insert:plasmid) molar ratio with the constructed *pUC19-pro-term* plasmid, which was also cut at the both *XhoI* and *BglII* restriction digest sites. T4 ligation was applied as described above, and the mixture was chemically transformed into *E. coli BL21* competent cells.

### 3.2.2.5. Construction of pUC19-*pro-term-GFP* and pUC19-*pro-term-mcherry*

Green fluorescent protein (*GFP*) and *mCherry* genes were obtained and amplified separately. Primers designed and used are also listed in Table 3.1. The PCR products were digested, respectively, with *NdeI/PstI* and *NdeI/ HindIII*. The PCR digested products were recovered for ligation with the pUC10-*pro-term* vector. The *pUC19-pro-term* plasmid was extracted from *E. coli Top10* cells and the *lacZ* gene was removed by digestion with either *NdeI/PstI* (for *GFP* insertion) or *NdeI/HindIII* (for *mCherry* insertion). The linearized vector was recovered by gel-band purification and mixed with digested *GFP* or *mCherry* PCR products. Ligation was implemented by T4 ligase at 16 ℃ overnight. The ligated vectors of *pUC19-pro-term-GFP* and

*pUC19-pro-term-mcherry* were then transformed into *E. coli Top10* cells and plated on solid LB medium with 100 μg/mL ampicillin and 100 μL X-gal/IPTC. Positive (white) colonies were selected and sub-cultured to harvest plasmid *pUC19-pro-term-GFP* and *pUC19-pro-term-mcherry*. Correct *GFP* and *mCherry* gene inserts were verified by PCR.

### 3.2.2.6. Construction of pUC19-pro-term-GFP-asGFP and pUC19-pro-term-mcherry-ascherry

The *GFP* and *mCherry* genes were also used as templates for PCR amplification of asRNA genes. These were referred to as *asGFP* and *asmCherry* DNA fragments. Primers used for this process are also listed in Table 3.1. The PCR reaction reagents and programs were the same as those described above. In all, 15 asGFP and 10 asmCherry DNA fragments were PCR amplified in single PCR reactions for each gene and cloned into the linearized vector containing the *GFP* or *mCherry* genes, respectively, using InFusion cloning (Clontech) according to the manufacturer's protocols. The InFusion cloning approach was used in this step, as opposed to the traditional "sticky-end" cloning used for *aslacZ* fragments because of significant difficulties obtaining high colony counts upon cloning with traditional methods and the *BglII* and *XhoI* restriction sites. These restriction sites were used for the InFusion approach, but this technique results in 15 bp overhangs. This was found to increase cloning efficiency dramatically and enabled all PCR fragments to be cloned simultaneously and located on a single plate. This led to a combinatorially designed experiment for evaluating asRNA genes.

To enable InFusion cloning, the pUC19-*pro-term-GFP* and pUC19-*pro-term-mcherry* vectors were harvested and digested overnight with *BglII/XhoI*. The digested product was purified and

treated with Antarctic phosphatase in the following reaction: (i) 20 µL digested vector (pUC19-pro-term-GFP/ pUC19-pro-term-mcherry) (10pg-10ng/ml), (ii) 5 µL 10X antarctic phosphatase buffer, (iii) 1 µL antarctic phosphatase, and (iv) 24 µL molecular biology grade water (total volume was 50 µL). The reaction was carried out at 37 ℃ for 2 hr, and the antarctic phosphatase was deactivated by incubation in a 65 ℃ water bath for 10 min. The deactivated antarctic phosphatase reaction solution containing digested vectors was used directly for InFusion cloning with the PCR purified products (*asGFP* and *asmCherry* DNA fragments). The InFusion cloning was performed according to the following procedure: (i) 2 µL mixtures of vector and insert both with the concentration of 10pg-10ng/ml were added to the reaction with the molar ratio of 1:3 (vector:insert), (ii) 2 µL 5X InFusion HD enzyme premix buffer was added, (iii) the total volume was adjusted to 10 µL with molecular biology grade water, (iv) the reaction was incubated at 50 ℃ for 15 min, (v) the reaction was then placed on ice for 30 min for chemical transformation into *E. coli 10beta* competent cells (Clontech). After transformation and growth on solid LB media overnight at 37 ℃, several colonies were selected and the asRNA gene insert was determined by PCR using primers designed for the *plac* promoter (left) and the *lacZ* terminator (right).

### 3.2.3. Determination of the targeting gene expression level

#### 3.2.3.1. *lacZ* gene expression level determination

The *lacZ* gene expression level was determined using a beta-galactosidase enzyme assay system kit. Beta-galactosidase is widely used as a gene reporter molecular, which is partially synthesized by *lacZ* gene. The beta-galactosidase enzyme assay system kit was able to assay beta-

galactosidase activity in cell lysates. The *E. coli BL21* host containing the *pUC19-pro-term-aslacZ* vector was cultured in LB medium with 100 μg/mL ampicillin at 37 ℃ for 6 hours. IPTG with a final concentration of 0.8 mM/L was added to induce the express of *lacZ* and *asLacZ* genes. The culture was incubated at 25 ℃ overnight. When the *E. coli BL21* culture was harvested, the cells were collected by centrifugation at 10,000 rpm for 10 min at 4 ℃. A 1XPBS buffer was applied to wash the cell pellet. After discarding the supernatant, the pelleted cells were re-suspended in 1XRLB buffer (Promega) at room temperature for 15 min to lyse the cells. The supernatant was collected, and the insoluble materials were removed by centrifugation. For each sample, 150 μL of cell lysate was mixed with the same amount of 2X beta-galactosidase enzyme assay buffer in a single well of a 96-well plate. After incubation for 30 min at 37 ℃, yellow color was observed and the intensity was determined at 420 nm by plate reader. After dividing by the concentration of the host cell (OD600), the results indicated the specific concentration of beta-galactosidase enzyme produced in each sample. This was used to interpret the *lacZ* gene expression level, which was affected by different *aslacZ* constructs.

#### 3.2.3.2. GFP and mCherry expression level determination

Since the strategy was to use a combinatorial design in cloning all *asGFP* and *asCherry* fragments simultaneously, the positive colonies on each plate were selected and sub-cultured into 24-well plate with the finial concentration of IPTG 0.8 mM/L and ampicillin 100 μg/mL. The culture was incubated in a plate reader at 25 ℃ with slow shaking. OD600 and fluorescence readings were determined every 30 min. For GFP the excitation wavelength of 480 nm was used, and the emission wavelength was 515 nm. For mCherry the excitation wavelength was 565nm, and the emission was 605nm. The large differences observed between the peak and bottom of the

fluorescence intensities indicated the *GFP* and *mCherry* gene expression levels were significantly affected by different *asGFP* and *asmCherry* genes.

Table.3.3. PCR primers designed and used in this research. The restriction digest sites are marked in italics and are underlined

| **Primer name** | **Part Amplified** | **Sequence** | **Restriction site** |
|---|---|---|---|
| pLacZ_left | *plac* promoter of pUC19 | AATGC*AGGTCCT*GTTGGCCGATTCATTA | *EcoO109I* |
| pLacZ_right | *plac* promoter of pUC19 | AGTTAG*CTCGAG*AATCTAAGTCAA*AGATCT*AATTCCACACAACATACG | *XhoI, BglII* |
| lacZ_term_left | *lacZ* terminator of pUC19 | ACTGTA*CTCGAG*TAAAGTTGAATCGTCAGTTT | *XhoI* |
| lacZ_term3_right | *lacZ* terminator of pUC19 | ATTAGA*GACGTC*AATTACCTGATGGACTGG | *AatII* |
| GFP_clone_left | *GFP* | ATTA*CTGCAG*ATGGTGAGCAAGGGCGA | *PstI* |
| GFP_clone_right | *GFP* | GTTG*CATATG*TTACTTGTACAGCTCGTCCATGC | *NdeI* |
| RFP_clone_left | *mCherry* | ATTA*AAGCTT*ATGGTGAGCAAGGGCGA | *HindIII* |
| RFP_clone_right | *mCherry* | GTTG*CATATG*TTACTTGTACAGCTCGTCCATG | *NdeI* |
| asLacZ1_left | *aslacZ genes* | ATGTAA*AGATCT*TTGTTAAAGTGTGTCCTTTG | *BglII* |
| asLacZ1_right_ | *aslacZ genes* | ATTTAC*CTCGAG*TACGTTCGAACCGC | *XhoI* |
| asLacZ3_left_ | *aslacZ genes* | ATGTAA*AGATCT*TCCTCTAGAGTCGACCTG | *BglII* |
| asLacZ2_right | *aslacZ genes* | ATTTAC*CTCGAG*ACACAGGAAACAGCTATGA | *XhoI* |
| asLacZ3_left_ | *aslacZ genes* | ATGTAA*AGATCT*CCAGCTGGCGAAA | *BglII* |
| asLacZ3_right | *aslacZ genes* | ATTTAC*CTCGAG*TGGGAAAACCCTGG | *XhoI* |
| asGFP_1 left | *asGFP* genes | TTCAACTTTA*CTCGAG*GTGAGCAAGGGCGAGGA | *XhoI* |
| asGFP_1 right | *asGFP* genes | GTGTGGAATT*AGATCT*GAACTTCAGGGTCAGCTTGC | *bglII* |
| asGFP_2 left | *asGFP* genes | TTCAACTTTA*CTCGAG*TGGCC | *XhoI* |

| | | | |
|---|---|---|---|
| | | CACCCTCGTG | |
| asGFP_2 right | *asGFP* genes | GTGTGGAATT*AGATCT*GGGTA GCGGGCGAAGC | *bglII* |
| asGFP_3 left | *asGFP* genes | TTCAACTTTA*CTCGAG*_AGCA GCACGACTTCTTCAAG | *XhoI* |
| asGFP_3 right | *asGFP* genes | GTGTGGAATT*AGATCT*AGAAG ATGGTGCGCTCCTG | *bglII* |
| asGFP_4 left | *asGFP* genes | TTCAACTTTA*CTCGAG*TGAAG TTCGAGGGCGACA | *XhoI* |
| asGFP_4 right | *asGFP* genes | GTGTGGAATT*AGATCT*_TGTGC CCCAGGATGTTG | *bglII* |
| asGFP_5 left | *asGFP* genes | TTCAACTTTA*CTCGAG*CTATAT CACCGCCGACAAGC | *XhoI* |
| asGFP_5 right | *asGFP* genes | GTGTGGAATT*AGATCT*GAACTC CAGCAGGACCATGT | *bglII* |
| asRFP_1left | *asmCherry* genes | TTCAACTTTA*CTCGAGGGCCAT CATCAAGGAGTTCA* | *XhoI* |
| asRFP_1right | *asmCherry* genes | GTGTGGAATT*AGATCT*ACATGA ACTGAGGGGACAGG | *bglII* |
| asRFP_2left | *asmCherry* genes | TTCAACTTTA*CTCGAGGACGG CGAGTTCATCTACAA* | *XhoI* |
| asRFP_2right | *asmCherry* genes | GTGTGGAATT*AGATCT*CCCATG GTCTTCTTCTGCAT | *bglII* |
| asRFP_3left | *asmCherry* genes | TTCAACTTTA*CTCGAG*CCTGA AGGGCGAGATCAAG | *XhoI* |
| asRFP_3right | *asmCherry* genes | GTGTGGAATT*AGATCT*TCTTGG CCTTGTAGGTGGTC | *bglII* |
| asRFP_4left | *asmCherry* genes | TTCAACTTTA*CTCGAG*ACATCA CCTCCCACAACGAG | *XhoI* |
| asRFP_4right | *asmCherry* genes | GTGTGGAATT*AGATCT*CTTGTA CAGCTCGTCCATGC | *bglII* |

## 3.3. Results

### 3.3.1. Results of pUC19-pro-term plasmid construction

The PCR amplification of *plac* promoter and *lacZ* terminator were visualized by 2% agarose gel and shown in Fig.3.4a. The size of *plac* promoter is 200 bp and the size of *lacZ* terminator is about 250 bp. To determine if cloning was successful (the ligation of promoter and terminator),

the constructed plasmid *puc19-pro-term* was used as the DNA template to amplify the insertion region with the primers of *pLacZ_left* and *lacZ_term_right* (see Table 3.3). The PCR result was expected to be 450 bp, and this result is shown correctly in Fig. 3.4b. The recombinant sequence of promoter and terminator was sequenced by the core facilities at the Virginia Bioinformatics Institute to further ensure the cloning was successful.

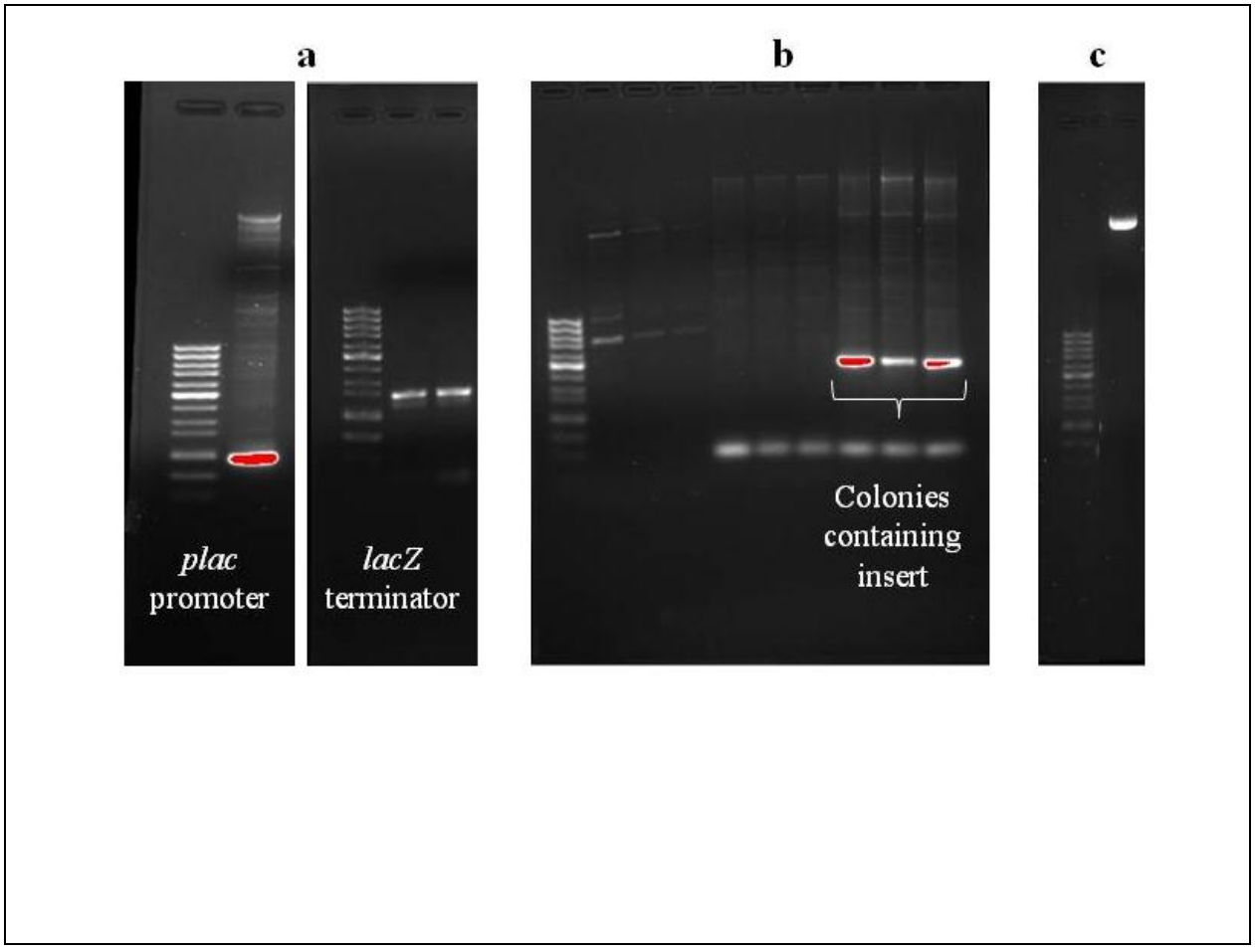



Fig. 3.4. Gel electrophoresis showing the following: (a) amplification of the *plac* promoter and *lacZ* terminator, (b) successful cloning of the cassette containing the promoter and terminator, and (c) uncut pUC19-*pro-term* plasmid.

### 3.3.2. Results of aslacZ gene fragment amplification

The *aslacZ* gene fragments were PCR amplified and visualized by 2% agarose gel. The gel results of 23 samples are shown in Fig. 3.5 and demonstrate the different sizes of fragments created by combinations of the forward and reverse primers. A 50 bp mini ladder was used to determine the fragments sizes.

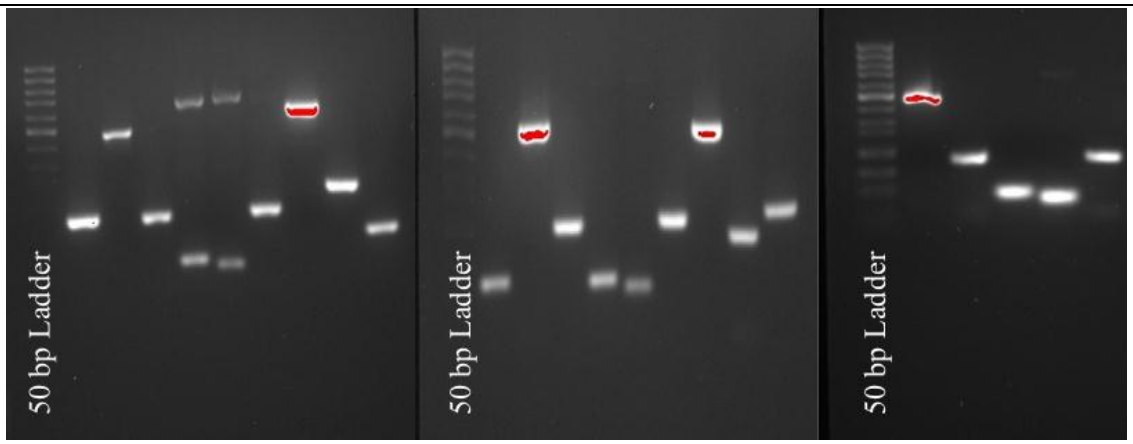



Fig. 3.5. A total of 23 samples of *aslacZ* DNA fragments. The following sample IDs are shown: aA, bA, bB, bC, bD, bE, cA, cB, cC, cD, dA, dB, dC, dD, dE, eA, eB, cE, eA, eB, eC, eD, eE.

### 3.3.3. Results of asGFP and asmCherry amplification and the construction of pUC19-pro-term-gfp and pUC19-pro-term-mChery plasmids

The pUC19-*pro-term* plasmid was digested and the *lacZ* gene was removed as described above. Fig.3.6 shows the PCR amplification of *GFP* gene from the pUC19-*pro-term-gfp* vector and the *mCherry* gene from pUC19-*pro-term-mCherry* plasmid. Both gel results displayed the proper

size inserts to suggest successful cloning. In addition, both plasmids displayed the proper fluorescent properties. This is discussed in detail in a later section.

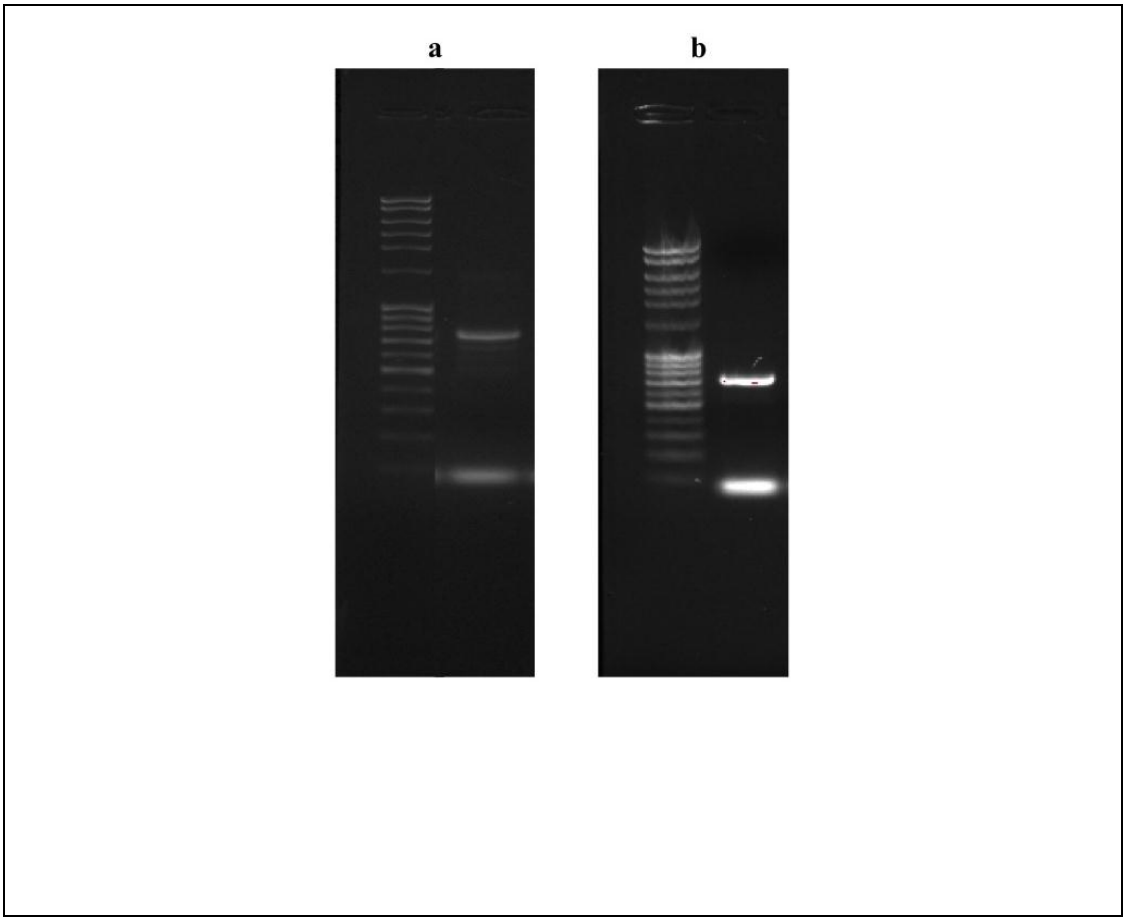



Fig. 3.6. PCR verification of successful cloning of: (a) *GFP* and (b) *mCherry* into the pUC19-*pro-term* vector.

### 3.3.4. Amplification of asGFP and asmCherry DNA fragments

The *asGFP* and *asmCherry* DNA fragments were separately PCR amplified and visualized by gel electrophoresis. Fig. 3.7 shows the PCR results of 15 randomly selected *asGFP* DNA fragments. Each fragment was obtained from a different colony growing on the same solid medium plate. Cloning efficiency was observed to be very high with the InFusion cloning

methods, and thousands of colonies were obtained per plate. This enabled an entire library of *asGFP* and *asmCherry* fragments to be cloned simultaneously. Fig. 3.8 shows 10 randomly selected *asmCherry* DNA fragments obtained from 10 different colonies of the plate.

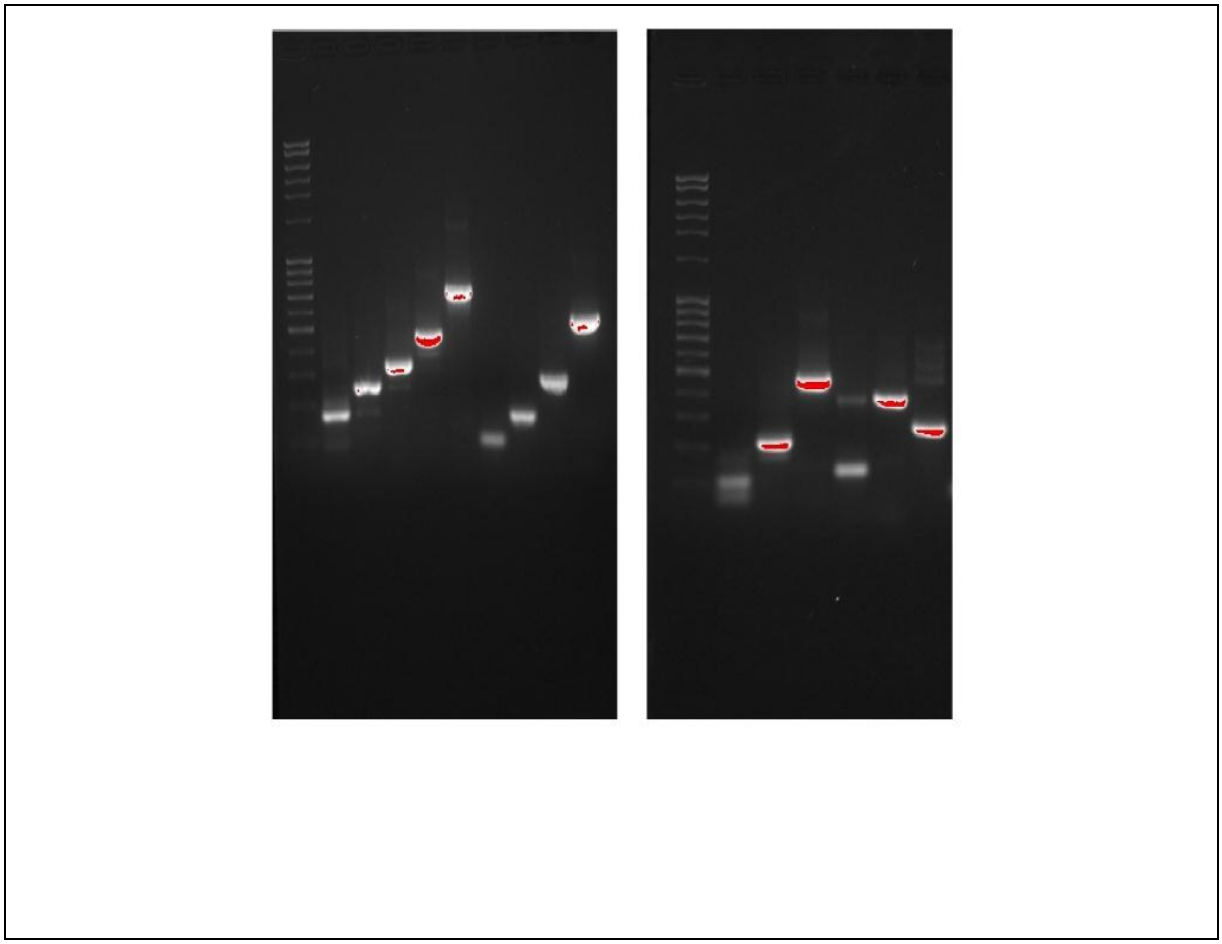


Fig. 3.7. PCR amplification of 15 *asGFP* DNA fragments with a mid-range ladder. The following sample IDs are shown: 11,12,13,14,15,22,23,24,25, 33,34,35,44,45,55.

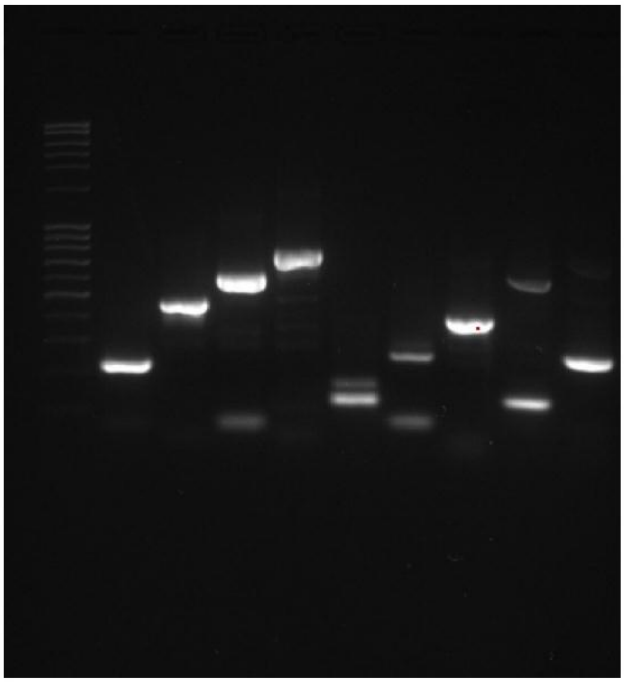


Fig. 3.8. PCR amplification results for 10 randomly selected asmChery DNA fragments with a mid-range ladder. The following sample IDs are shown: 11,12,13,14,22,23,24, 33, 34, 44.

### 3.3.5. Identification of asGFP and asmCherry fragments

The positive *E. coli 10beta* colonies with pUC19-*pro-term-gfp-asgfp* or pUC19-*pro-term-mCherry-asmCherry* plasmids were sub-cultured in a single well of a 24-well plate and their fluorescence intensities were monitored (separately) with growth in the plate reader incubator at 25 ℃ for 45 hours. The results of green and cherry fluorescence for several different colonies and asRNA genes are listed in Table 3.4. The identity of the asRNA gene was determined by PCR with the *pLacZ_left* and *lacZ_term_right* primers (see Table 3.3). The PCR products included the *plac* terminator, the asRNA gene, and the *lacZ* terminator. As stated previously, the promoter and terminator had a combined size of 450 bp. Selected examples of asRNA fragment

identification are shown in Fig. 3.9. In most cases, size was sufficient to identify the asRNA gene fragment. However, in certain cases, fragments of comparable sizes existed. In this case, it was necessary to perform additional PCR runs using primers of the candidate fragments.

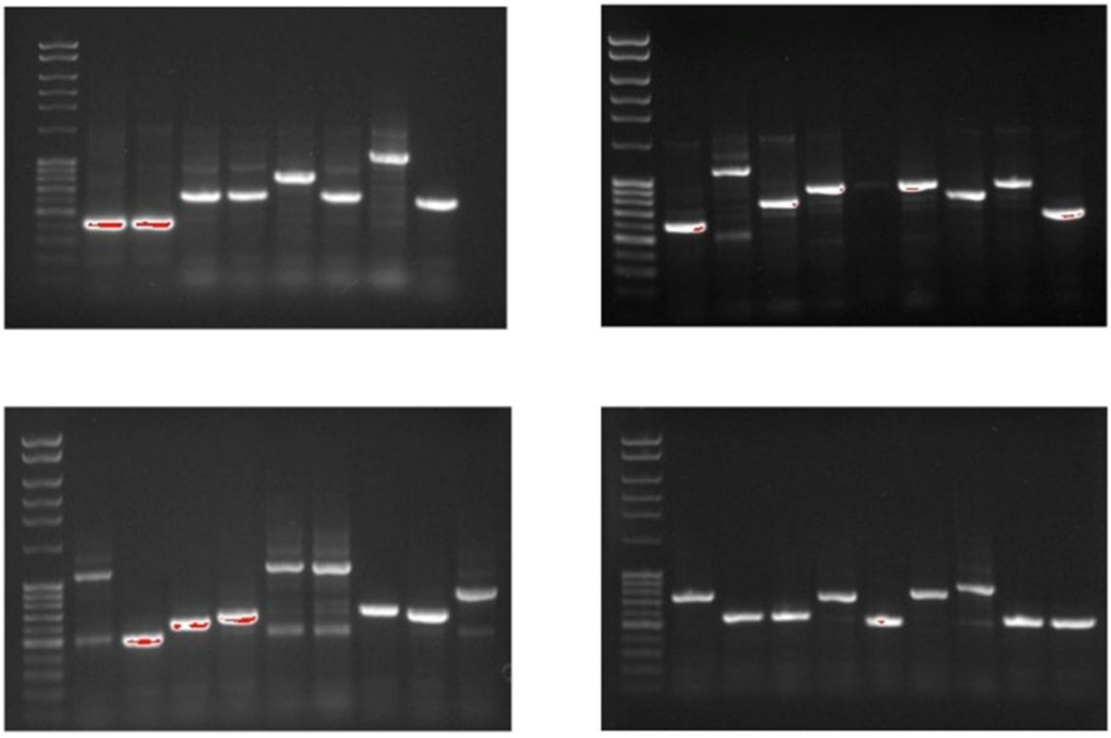


Fig. 3.9. PCR identification of asRNA fragments. The insertion results of pro-asgfp-term in GFP expression E. coli 10beta are shown in the top row and those of pro-asmCherry-term are shown in the bottom two panes.

### 3.3.6. Quantitative analysis of lacZ expression level

The *lacZ* gene and *aslacZ* DNA fragments were all installed under the control of individual copies of the *plac* promoter. Thus, the same promoter/terminator combination was expected to yield identical levels of transcripts for the *lacZ* and *asLacZ* genes. The hybridization of transcribed *aslacZ* DNA fragments with mRNA of *lacZ* resulted in the down-regulation of *lacZ*

translation in most cases and up-regulation was observed for some cases as well. The *lacZ* gene expression level was determined by the beta-galactosidase enzyme assay kit. Values of the absorption intensity at 420 nm divided by that of OD600 were recorded. The absorption intensity at 420 nm provided the concentration information of beta-galactosidase, while the OD600 provided the cell density. The measured numerical values and calculated thermodynamic values for each asRNA fragment using NUPACK are listed in Table 3.4. Results are also shown graphically as Fig. 3.10. Here, it is also shown that the *lacZ* gene expression level is impacted by *aslacZ* DNA fragments upon comparison with the control. The results of thermodynamic analyses are also given in Table 3.4. The binding percentage for each fragment with *lacZ* mRNA is given along with the minimum free energy (kcal/mol) of the hybridized complex. In several instances, the expression level increases in the presence of an asRNA. There is no known instance where an asRNA results in increased translation of that mRNA. While this is an extremely interesting concept, the extreme amount of experimental error observed with this data calls into question the validity of this result. Correlations were sought between the observed expression level of the *lacZ* gene and the calculated thermodynamic parameters (i) binding percentage with *lacZ* mRNA and (ii) minimum free energy. The binding percentage parameter identifies the amount of *aslacZ* hybridizing with *lacZ* mRNA, and the minimum free energy parameter describes the Gibbs free energy of the hybridized asRNA-mRNA complex. Several relationships are plotted in Fig. 3.11. As shown, no significant relationships exist between the expression level and either the binding percentage or the minimum free energy for the *lacZ* gene. Correlation coefficients were calculated to quantify the relationship. A weak relationship exists between the binding percentage and minimum free energy, as shown in Fig. 3.11C. Because of the lack of relationships observed for *lacZ* expression in the presence of asRNA fragments, these

experiments were repeated using the *GFP* and *mCherry* genes. The enzymatic assay was thought to provide significant experimental error to the system, making *lacZ* a poor choice for this quantitative analysis.

Table 3.4. Determination of *lacZ* expression level and thermodynamic parameters

| Sample ID | Absorbance 420 nm | Standard Deviation | OD600 | Expression level compared with control | Binding percentage with *lacZ* mRNA | Minimum free energy (kcal/mol) |
|---|---|---|---|---|---|---|
| Control | 0.483667 | 0.040526 | 0.213 | 1.00 | N/A | N/A |
| aA | 0.429 | 0.07015 | 0.238 | 0.793805 | 96.87 | -184.08 |
| bA | 0.398333 | 0.057813 | 0.302 | 0.580862 | 42.99 | -233.88 |
| bB | 0.253667 | 0.043004 | 0.265 | 0.421552 | 10.07 | -168.98 |
| bC | 0.234667 | 0.022942 | 0.259 | 0.399011 | 1.26 | -143.78 |
| bD | 0.423 | 0.097411 | 0.377 | 0.49412 | N/A | N/A |
| bE | 0.256667 | 0.03499 | 0.259 | 0.436418 | 10.2 | -169.48 |
| cA | 0.435 | 0.111853 | 0.261 | 0.733977 | 49.38 | -247.38 |
| cC | 0.662333 | 0.246151 | 0.231 | 1.262694 | 5.1 | -155.98 |
| cD | 0.377333 | 0.10232 | 0.288 | 0.576987 | N/A | N/A |
| cE | 1.034333 | 0.165868 | 0.248 | 1.836717 | 1.1 | -182.98 |
| dA | 1.02 | 0.017349 | 0.239 | 1.879471 | 48.94 | -242.88 |
| dB | 0.242667 | 0.027392 | 0.269 | 0.397275 | 13.19 | -177.98 |
| dC | 0.619 | 0.073912 | 0.238 | 1.145374 | 0.8746 | -152.78 |
| dE | 0.366667 | 0.026577 | 0.281 | 0.574644 | 11.67 | -178.48 |
| eA | 0.783667 | 0.050935 | 0.204 | 1.691744 | 41.8 | -238.68 |
| eB | 0.390667 | 0.002517 | 0.242 | 0.710926 | 8.732 | -173.78 |
| eC | 1.014333 | 0.050402 | 0.216 | 2.068047 | 0.439 | -147.08 |
| eD | 0.289333 | 0.029956 | 0.46 | 0.276996 | N/A | N/A |
| eE | 0.485667 | 0.054455 | 0.313 | 0.683325 | 8.96 | -147.28 |

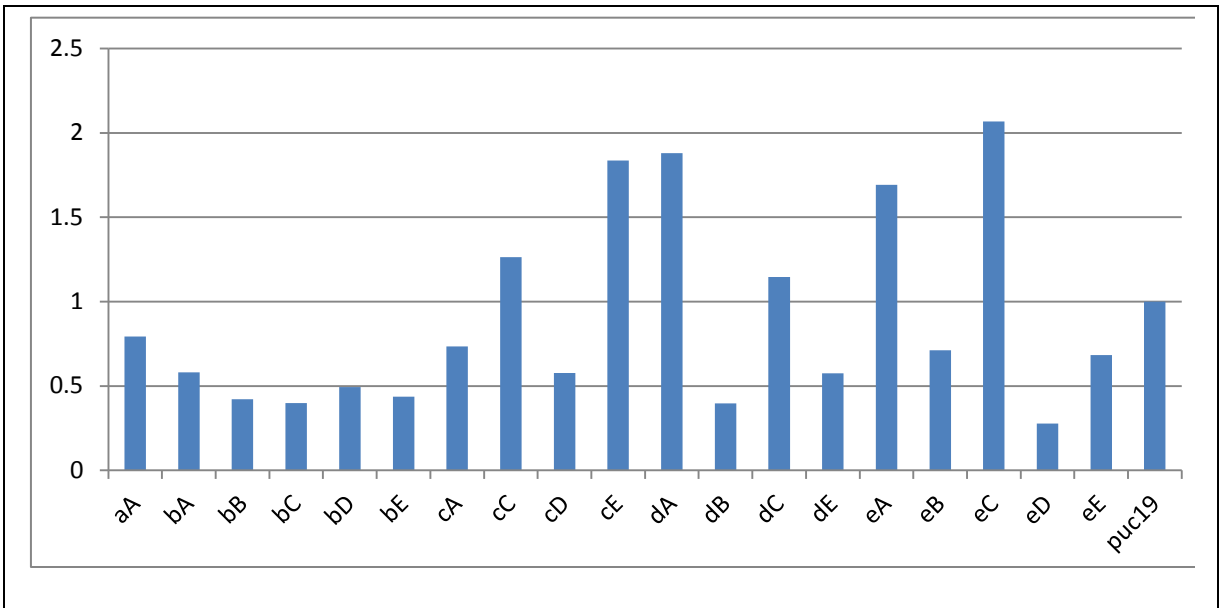



Fig. 3.10. The *lacZ* gene expression level in the presence of specified *asLacZ* fragments.

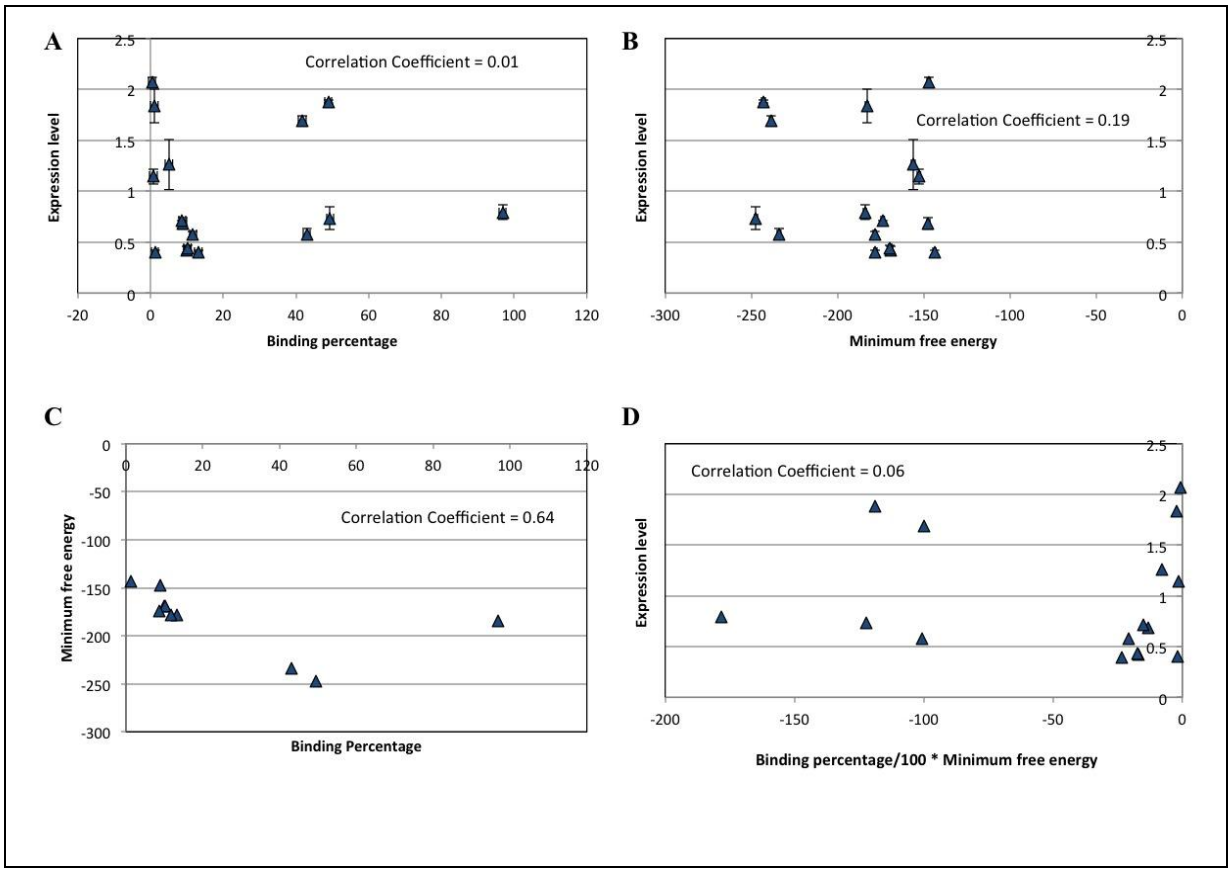



Fig. 3.11. Relationships between the observed expression level of *lacZ* in the presence of an *aslacZ* fragment (relative to the control) as a function of: (A) binding percentage calculations and (B) minimum free energy of the hybridized complex. The relationship between binding percentage and minimum free energy is shown in (C), and the expression level as a function of both binding percentage and minimum free energy is shown in (D).

### 3.3.7. Analysis of GFP and mCherry expression levels

The *GFP* and *mCherry* genes are the superior choice for reporters, compared to *lacZ*, because gene expression levels could be analyzed along with culture growth in real time using a plate reader. Thus, *GFP* and *mCherry* gene expression levels were analyzed, respectively, for several *aslacZ* fragments simultaneously using the plate reader and the fluorescence excitation and emission wavelengths of 480/515 for GFP and 565/605 for mCherry. As the cloning strategy was combinatorial in nature to obtain positive colonies, some of the sequences were not found on the

plate. The expression levels are listed in Table 3.5 with other crucial thermodynamic indicators, including the binding percentage and the minimum free energy. The relationships between thermodynamic parameters and the expression level (relative to the control with no asRNA fragment) are plotted in Fig. 3.12. With the fluorescent reporter genes strong correlation coefficients were observed between the binding percentage and the expression level. These exceeded 0.9 for the *GFP* and *mCherry* genes individually, and the correlation coefficient was 0.72 when both genes were considered together. Weak correlation was observed between the expression level and the minimum free energy parameter. However, when the binding percentage and the minimum free energy were combined (Fig. 3.12D), a superior correlation coefficient (0.79) was obtained for the total data set. This suggests that both the binding percentage and the minimum free energy should be used in designing "fine-tuned" asRNA fragments.

Table 3.5. Expression level of *GFP* gene (above) and *mCherry* gene (below) in the presence of asRNA gene fragments.

| Sample ID (*GFP*) | Expression level | Reduced expression level | Binding percentage | Minimum free energy (kcal/mol) |
|---|---|---|---|---|
| 1_1 | 0.77 | 0.23 | 98.35 | -283.58 |
| 1_2 | 1.04 | -0.04 | 6.849 | -0.58403 |
| 1_3 | 0.765 | 0.235 | 96.83 | -328.88 |
| 1_4 | 0.93 | 0.07 | 32.72 | -368.18 |
| 1_5 | 0.935 | 0.065 | 35.49 | -428.68 |
| 2_2 | 0.74 | 0.26 | 96.21 | -242.88 |
| 2_3 | 0.77 | 0.23 | 97.83 | -259.48 |
| 2_4 | 0.84 | 0.16 | 42.18 | -298.38 |
| 2_5 | 0.835 | 0.165 | 43.1 | -359.28 |
| 3_3 | 0.735 | 0.265 | 95.37 | -243.48 |
| 3_4 | N/A | N/A | 74.14 | -284.28 |
| 4_4 | 0.81 | 0.19 | 78.16 | -254.58 |
| 4_5 | 0.9 | 0.1 | 47.53 | -314.98 |
| 5_5 | 0.84 | 0.16 | 42.74 | -279.98 |

| Sample ID (*mCherry*) | Expression level | Reduced expression level | Binding percentage | Minimum free energy (kcal/mol) |
|---|---|---|---|---|
| 1_1 | 1.02 | -0.02 | 0.9 | -322.98 |
| 1_2 | 0.943 | 0.057 | 8.64 | -403.08 |
| 1_3 | 0.745 | 0.255 | 37.32 | -441.38 |
| 1_4 | 0.65 | 0.35 | 48.36 | -485.78 |
| 2_2 | 0.54 | 0.46 | 98.25 | -271.18 |
| 2_3 | 0.74 | 0.26 | 43.73 | -309.48 |
| 2_4 | 0.685 | 0.31 | 47.78 | -353.88 |
| 3_3 | 0.99 | 0.01 | 2.047 | -264.48 |
| 3_4 | 0.86 | 0.14 | 24.18 | -309.28 |
| 4_4 | 0.97 | 0.03 | 5.921 | -275.58 |

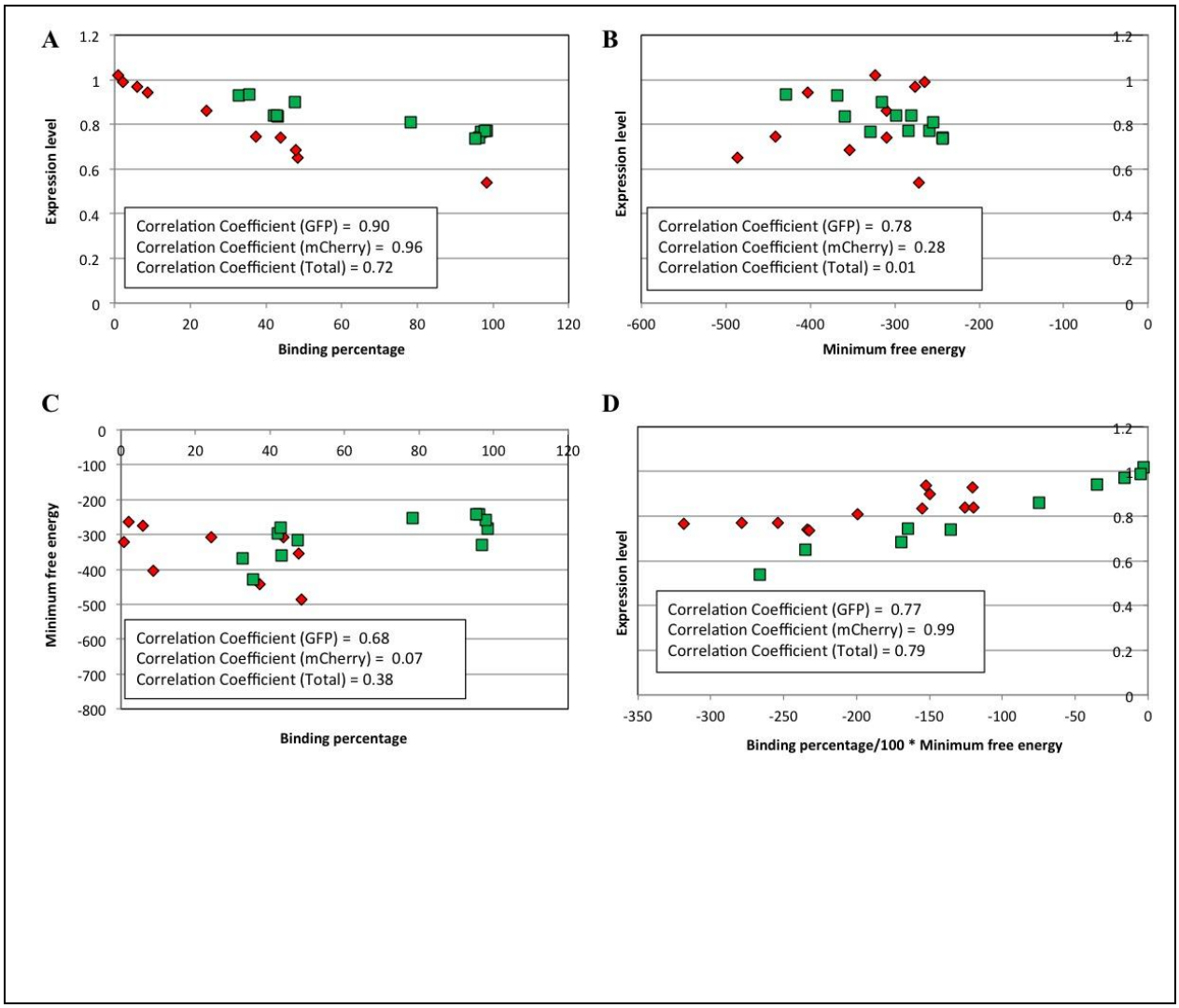
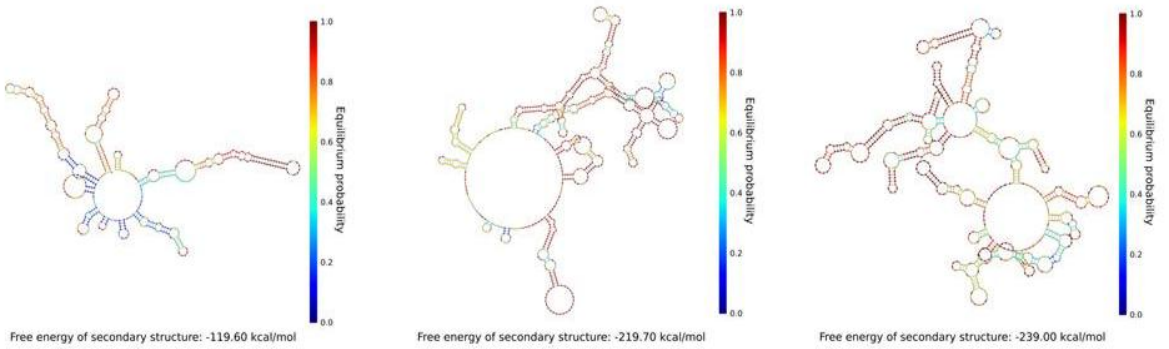



Fig. 3.12. Relationships between the observed expression level of *GFP* (green) and *mCherry* (red) genes in the presence of asRNA fragments (relative to the control) as a function of: (A) binding percentage calculations and (B) minimum free energy of the hybridized complex. The relationship between binding percentage and minimum free energy is shown in (C), and the expression level as a function of both binding percentage and minimum free energy is shown in (D).

### 3.3.8. Secondary structure predictions of lacZ, GFP and mCherry mRNA

The *lacZ*, *GFP,* and *mCherry* genes were first transcribed into mRNA, the combination of asRNA was operated based on the formed secondary structure of mRNA. The RNA folding program mfold was used to predict the secondary structure of these mRNA molecules. They are shown in Fig. 3.13.




A. *lacZ* mRNA B. *GFP* mRNA C. *mCherry* mRNA

Fig. 3.13. Predicted secondary structures of: (A) lacZ mRNA, (B) GFP mRNA, and (C) mCherry mRNA. Predictions were generated using the mfold program (ref).

### 3.3.9. Hybridization of asRNA fragments with mRNA

The secondary structure of all potential *aslacZ-lacZ* mRNA complexes, binding percentage of asRNA fragments with the target mRNA, and minimum free energy were predicted using mfold and NUPACK software. As an example, the complex of *aslacZ* RNA (Sample ID: aA) with the *lacZ* mRNA is showed in Fig. 3.14. In addition, example complexes between *asGFP* and *GFP* mRNA and *asmCherry* and *mCherry* mRNA are given in Fig. 3.15.

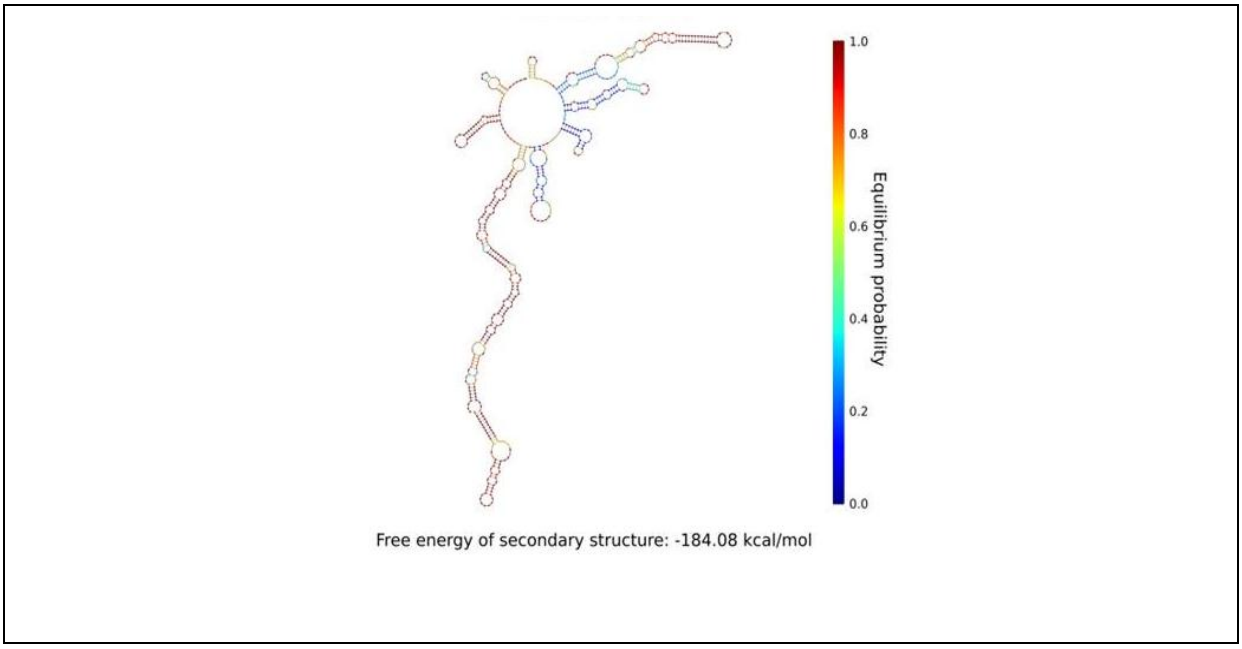



Fig. 3.14. The complex of *aslacZ* (Sample ID: aA) with *lacZ* mRNA as predicted by mfold.

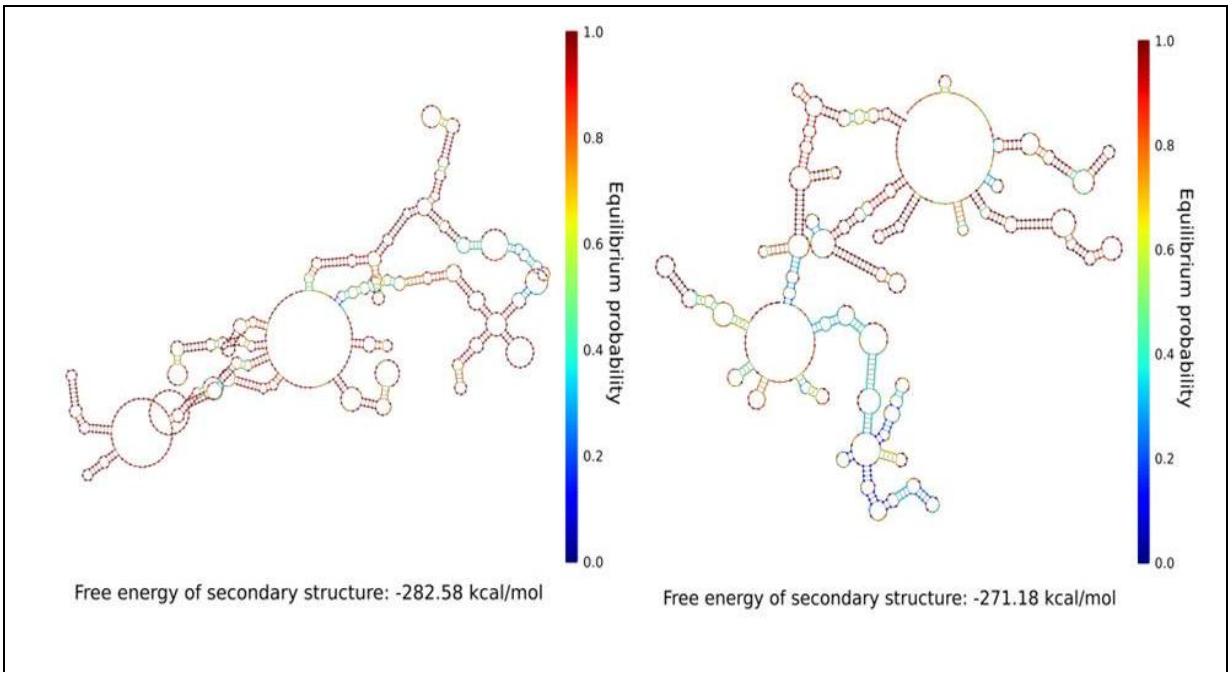



Fig. 3.15. The secondary structure of the *asGFP* with *GFP* mRNA complex (left, sequence ID 1_1) and secondary structure of the as*mCherry* with *mCherry* mRNA complex (right, sequence ID 2_2).

## 3.4. Discussion

### 3.4.1. Antisense RNA regulation

The asRNA genes inhibited the translation of targeted gene mRNA through hybridization following transcription. Although different constructs showed a wide variety in their capabilities for down-regulation, the overall approach involved the synthesis of a short transcript RNA that does not code for protein but is complementary with a target mRNA molecule, enabling the two to hybridize [28]. Many studies have been published on this topic to describe the phenomena of naturally occurring asRNA in prokaryotes [28]. These instances provide the metabolic engineer an additional tool to the use for the manipulation of cellular function. For example, it is often desired to redirect cellular metabolism towards a particular bioproduct. This can be done by gene over-expression and competing pathway knockout, but more elegant strategies may involve the partial shut-down of a pathway, instead of its complete knockout. Moreover, this strategy in designing artificial asRNA genes for "fine-tuning" gene expression was even proposed and applied in some biopharmaceutical applications [2]. In recent years, increasing attention has been paid to the use of antisense approaches to engineer alterations in cellular pathways, including metabolism pathway regulation by altering crucial gene expression and metabolic fluxes. For instance, Bailey and coworkers have successfully used dual-regulated control vectors to investigate critical cell-cycle regulatory pathways. [31]

### 3.4.2. Problems on antisense RNA application in prokaryotes

However, because asRNA techniques involved complementary base pairing to a target mRNA, they were, in principle, very specific. Delivery, regulation, cellular specificity and other various questions regarding asRNA combination with targeting mRNA were all required to be addressed comprehensively [29]. Although, advanced software greatly benefited on the simulation of

mRNA secondary structure and the prediction of the interaction of two RNA hybridizations, the exact level of suppression of translation of the targeted mRNA were unknown until now, and the factors that affected asRNA effectiveness remained of research interest. Three model systems were examined in this research: (i) asRNA regulating the *lacZ* gene, (ii) asRNA regulating *GFP*, and asRNA regulating the *mCherry* fluorescent reporter gene. These were studied to explore the relationships between expression level and the thermodynamics of asRNA hybridization. After realizing the thermodynamic relationships, the approach could be further applied to predict the expression level of targeting gene under any specific asRNA and regulate the metabolic pathway for special bioproducts through controlling the expression level of crucial enzymes. To achieve this overall goal of designing asRNA to "fine-tune" gene expression, solutions to two problems were required. First, how many asRNA molecules have the potential to hybridize with the target mRNA? Second, what is the result of hybridization? This means, are these complexes irreversible? What is the secondary structure of the complex? Is hybridization limited by reaction kinetics or thermodynamics?

### 3.4.3. Analysis of hybridization complex factors

To answer the first question, the software nucleic acid package (NUPACK) [15] offered a powerful function to calculate the binding percentage of asRNA and targeting mRNA under a given condition. The concentrations of asRNA and target mRNA as well as the reaction temperature were required. In this study, the asRNA and target mRNA were both under the control of the same *plac* promoter; thus, it was assumed that the transcriptional efficiencies of both the asRNA gene and the target gene were the same. The incubation temperature of the *E. coli* culture (25°C) was applied when calculating the combination percentage using the

NUPACK software. As to the second question, in previous studies [3], the GC content, the length of asRNA, the oligonucleotides sequence of loops on targeting mRNA and some other factors were explored. It was generally considered that the minimum free energy for hybridization was the most significant factor to impact the stability of asRNA-mRNA complex and further hinder translation. To calculate the minimum free energy of the complex, it was necessary to clarify the secondary structure of the complex. Several bioinformatics tools provide this information. The NUPACK software [15] and the program mfold [16] were used in this research. However, it was found that, overall, there is a weak correlation between the minimum free energy of the asRNA-mRNA complex and target gene expression level. Much greater influence was found from percentage of asRNA molecules participating in binding the mRNA.

### 3.4.4. Conclusions of the thermodynamic model

In this study, gene regulation as a result of targeting mRNA translation by asRNA was explored quantitatively. The model system consisting of combinatorially derived fragments included *aslacz, asgfp* and *asmCherry*, and resulting gene expression data were related to thermodynamic calculations. A strong correlation and near linear relationship was observed between the percentage of asRNA binding mRNA and the observed expression level for *GFP* and *mCherry* fluorescent reporter genes. No correlations were detected for *lacZ*, which required an enzymatic assay to quantify expression. Significant influence of the minimum free energy for the formation of the asRNA-mRNA complex was observed for the fluorescent reporters. Further study is required to improve these relationships and build a comprehensive model for the design of novel asRNA molecules to target specific genes in prokaryotes. Ultimately, the goal of this research is to produce model-designed asRNA fragments that can knockdown gene translation to a specified

level (e.g., 25% of wild-type levels). This will provide a new tool for metabolic engineering that will help redirect metabolic flux in cases where gene knockouts are not feasible or warranted. In addition, since asRNA genes are controlled by a promoter, an inducible promoter can be installed to produce a metabolic "switch" capable of turning pathways on/off through the addition of an inducer.

**3.5. References**

1. Zamecnik PC, Stephenson ML: Inhibition of Rous sarcoma virus replication and cell transformation by a specific oligodeoxynucleotide. Proc Natl Acad Sci USA 1978, 75:280-284.

2. Opalinska JB, Gewirtz AM: Nucleic-acid therapeutics: basic principles and recent applications. Nat Rev Drug Discov 2002, 1:503-514.

3. Li Kim Lee and Charles M Roth. Antisense technology in molecular and cellular bioengineering. Current Opinion in Biotechnology 2003, 14:505–511

4. Roth CM, Yarmush ML: Nucleic acid biotechnology. Annu Rev

Biomed Eng 1999, 1:265-297.

5.Seshu B. Tummala, Stefan G. Junne, and Eleftherios T. Papoutsakis. Antisense RNA Downregulation of Coenzyme A Transferase Combined with Alcohol-Aldehyde Dehydrogenase Overexpression Leads to Predominantly Alcohologenic *Clostridium acetobutylicum* Fermentations. J. Bacteriol. 2003, 185(12):3644

6. Ferrari J, Gunson J, Lofgren J, Krummen L, Warner TG: Chinese hamster ovary cells with constitutively expressed sialidase antisense RNA produce recombinant DNase in batch culture with increased sialic acid. Biotechnol Bioeng 1998, 60:589-595.

7. Dellambra E, Golisano O, Bondanza S, Siviero E, Lacal P,Molinari M, D'Atri S, De Luca M: Downregulation of 14-3-3r prevents clonal evolution and leads to immortalization of primary human keratinocytes. J Cell Biol 2000, 149:1117-1130.

8. Abogadie FC, Bron R, Marsh SJ, Drew LJ, Haley JE, Buckley NJ,

Brown DA, Delmas P: Adenovirus-mediated Gaq-protein antisense transfer in neurons replicates Gaq gene knockout strategies. Neuropharmacol 2002, 42:950-957.

9. Dargelos E, Moyen C, Dedieu S, Veschambre P, Poussard S,Vuillier-Devillers K, Brustis JJ, Cottin P: Development of an inducible system to assess p94 (CAPN3) function in cultured muscle cells. J Biotechnol 2002, 96:271-279.

10. Tummala SB, Welker NE, Papoutsakis ET: Design of antisense

RNA constructs for downregulation of the acetone formation pathway of Clostridium acetobutylicum. J Bacteriol 2003, 185:1923-1934.

11 Andrew B.Cubitt, Roger Heim, Stephen R. Adams. Understanding, improving and using green fluorescent proteins. Techniques.1995. Nov 448-455.

12. Peter van Roessel and Andrea H. Brand. Imaging into the future: visualizing gene expression and protein interactions with fluorescent proteins. NATURE CELL BIOLOGY. VOL 4 JANUARY 2002

13. Xiaokun Shu, Nathan C. Shaner, Corinne A. Yarbrough, Roger Y. Tsien, and S. James Remington. Novel Chromophores and Buried Charges Control Color in mFruits. Biochemistry, Vol. 45, No. 32, 2006

14. Daniel Yarbrough, Rebekka M. Wachter, Karen Kallio, Mikhail V. Matz, and S. James Remington. Refined crystal structure of DsRed, a red fluorescent protein from coral, at 2.0-Å resolution. PNAS. January 16, 2001 vol. 98, no. 2, 463

15. David H Mathews and Douglas H Turner. Prediction of RNA secondary structure by free energy minimization. Current Opinion in Structural Biology 2006, 16:270–278

16. Mathews DH, Zuker M. Predictive methods using RNA sequences in Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins, 2004:143-170.

17. Eddy SR: How do RNA folding algorithms work? Nat Biotechnol. 2004, 22:1457-1458.

18. Dirks, R. M.; Bois, J. S.; Schaeffer, J. M.; Winfree, E.; Pierce, N. A., SIAM Rev 2007, 49, 65.

19. Dirks, R. M.; Pierce, N. A. J Comput Chem 2003, 24, 1664.

20. Dirks, R. M.; Pierce, N. A. J Comput Chem 2004, 25, 1295.

21. Zeiler, B. N. & Simons, R. W. (1997). Antisense RNA structure and function. In RNA Structure and Function pp. 437-464, Cold Spring Harbor Laboratory Press,Cold Spring Harbor, NY.

22. Thomas Franch, Michael Petersen, E.Gerhart H Wagner, Jens Peter Jacobsen, Kenn Gerdes. Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure. Journal of Molecular Biology Volume 294, Issue 5, Pages 1115-1125

23. S. Brantl, Antisense-RNA regulation and RNA interference. Biochim Biophys Acta, 1575 (2002), pp. 15–25.

24. Robert M. Dirks, Justin S. Bois, Joseph M. Schaeffer, Erik Winfree, Niles A. Pierce. Thermodynamic Analysis of Interacting Nucleic Acid Strands. SIAM REVIEW Vol. 49, No. 1, pp. 65–88

25. D.H. Mathews, J. Sabina, M. Zuker, and D.H. Turner, Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure, J. Mol. Biol.,288 (1999), pp. 911–940.

26. J. SantaLucia, Jr., A unified view of polymer, dumbbell, and oligonucleotide DNA nearest neighbor thermodynamics, Proc. Natl. Acad. Sci. USA, 95 (1998), pp. 1460–1465.

27. I. Tinoco, Jr., O.C. Uhlenbeck, and M.D. Levine, Estimation of secondary structure in ribonucleic acids, Nature, 230 (1971), pp. 362–367.

28. Ophry Pines and Masayori Inouye. Antisense RNA regulation in Prokaryotes. TIG. November 1986. 284-287.

29. Masayori Inouye. Antisense RNA: its functions and applications in gene regulation - a review. Gene, 72 (1988) 25-34

30. Li Kim Lee and Charles M Roth. Antisense technology in molecular and cellular bioengineering. Current Opinion in Biotechnology 2003, 14:505–511

31. Moser S, Rimann M, Fux C, Schlatter S, Bailey JE, Fussenegger M: Dual-regulated expression technology: a new era in the adjustment of heterologous gene expression in mammalian cells. J Gene Med 2001, 3:529-549.

## 4. Conclusions and recommendations

### 4.1. Conclusions

In this research, the effective electroporation conditions of the fungus *P. irregulare* were determined for the first time. Electric shock of 2.0 kV with a 0.2cm cuvette was found to be the most effective condition for heterologous gene transformation. The auxotrophic selectable vectors (pESC-HIS, pESC-TRP, and pESC-URA) were transformed, respectively, into *P. irregulare* and cultured in their auxotrophic selection media. The yield and content of EPA and other components of total fatty acids (TFA) were further determined by the FAME approach using GC, as well as the analysis of biomass. The EPA content in *P. irregulare* with heterologous *TRP* gene reached 16.68 mg/g if cultured in tryptophan auxotrophic medium. This represented a 52.33% increase in EPA compared to the wild-type *P. irregulare*. The maximum of EPA yield was 98.52 mg/L from *P. irregulare* containing *URA* gene, a 32.28% increase over the wild-type. However, the maximum cell dried weight of these two organisms was, respectively, 6.13 g/L and 5.3 g/L. This was less than that of 6.80 g/L of the wild-type. Not only was a feasible approach detected to increase the EPA yield, this study also inferred that the ω-6 route was mainly involved in the EPA biosynthesis in *P. irregulare*.

Research was also performed on designing asRNA fragments by thermodynamics. Three model systems involving multiple asRNA fragments to the *lacZ* gene and the fluorescent reporter *GFP* and *mCherry* genes were constructed. These systems allowed the exploration of the quantitative relationships between the target gene expression level and designed asRNA sequence. This research also indicates that the percentage of asRNA binding mRNA is a crucial thermodynamic design parameter. The minimum free energy of hybridization between the asRNA and mRNA contributes to the expression level of the target gene, but the correlation is much weaker. The

binding percentage multiplied by the minimum free energy of hybridization was found to be a suitable model for predicting the expression level of a target gene in the presence of an asRNA. This thermodynamic model provides the possibility to predict the expression level of a target gene by any designed asRNA. This model can now be used in a design mode to create new asRNA molecules that can knockdown gene expression of a target by any degree desired.

### 4.2. Recommendations

To make the asRNA thermodynamic model to be more accurate, more samples are required and more complex thermodynamic models should be explored. Moreover, additional work is needed to ensure the model for a target gene can be adopted for another gene. The model must be universal to be applied broadly. This study also showed the linear relationships that exist between the binding percentage and expression level are limited. The inclusion of minimum free energy is required, but these relationships remain ambiguous. Additional data with more fluorescent reporters is required to explore this in detail. To implement the regulation by asRNA on EPA production in *P. irregulare* is still a long way away; although, the Δ5 desaturase, Δ6 desaturase and Δ6 elongases were considered as the crucial enzymes in EPA synthesis route. It will be suitable if overexpression or knockdown of the coding genes of these enzymes is investigated by a metabolic engineering strategy. However, in silico models are needed for *P. irregulare* to help determine how to fine tune metabolism. This approach might also benefit the EPA yield with the combination of optimizing the fermentation conditions.