CHARACTERIZATION AND EXPRESSION OF RECOMBINANT MOUSE GAROUPA (*Cromileptes altivelis*) GROWTH HORMONE AND INSULIN-LIKE GROWTH FACTOR GENES

AG. MUHAMMAD SAGAF B. ABU BAKAR

BIOTECHNOLOGY RESEARCH INSTITUTE
UNIVERSITI MALAYSIA SABAH
2007

UNIVERSITI MALAYSIA SABAH

BORANG PENGESAHAN STATUS TESIS@

JUDUL: Characterization and Expression of Recombinant Mouse Garoupa (*Cromileptes altivelis*) Growth Hormone and Insulin-like Growth Factor Genes

SARJANA: Sarjana Sains (Bioteknologi)

SESI PENGAJIAN: 2002 - 2007

Saya, AG. MUHAMMAD SAGAF B. ABU BAKAR mengaku membenarkan tesis Sarjana ini disimpan di Perpustakaan Universiti Malaysia dengan syarat-syarat kegunaan seperti berikut:

1. Tesis adalah hakmilik Universiti Malaysia Sabah.
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. TIDAK TERHAD

Disahkan oleh

ANITA BINTI ARSAD
PUSTAKAWAN KANAN
UNIVERSITI MALAYSIA SABAH

(Penulis: AG.MUHAMMAD SAGAF B. ABU BAKAR)

(TANDATANGAN PUSTAKAWAN)

Alamat Tetap:

(Penyelia: DR. CLEMENTE MICHAEL WONG VUI LING)

Tarikh: 5 April 2007

Tarikh: 22 Jun 2007

CATATAN: @ tesis dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah dan Sarjana secara penyelidikan, atau disertasi bagi pengajian secara kerja kursus dan penyelidikan, atau Laporan Projek Sarjana Muda (LPSM).

DECLARATION

The materials in this thesis are original except for quotations, excerpts, summaries and references, which have been duly acknowledged.

Ag. Muhammad Sagaf B. Abu Bakar

PS2002-013-511

5 April 2007

ACKNOWLEDGEMENT

In the name of God, the Merciful, the Compassionate.

I would like to take this opportunity to express my heartiest gratitude to my supervisor Dr. Clemente Michael Wong Vui Ling for helping me in my writing, and providing comments to further improve my thesis and co-supervisor, Prof. Dr. Saleem Mustapha for his suggestions, views and immediate attentions whenever I come to meet him. Most of all, I would like to thank Prof. Datin Dr. Ann Anton, Director of IPB, for her huge support and also the initiator of the project as one of the start-up projects for the institute.

I would also like to acknowledge Dr. Lee Ping Chin and her students for sharing their facilities, for me to conduct my recombinant protein research work, which was the detrimental critical bottle neck into achieving the final outcome.

To all the lecturers and staff of IPB, lab assistants, SST postgraduate and biotechnology research lab personnel and all my MSc. friends in IPB and SST which I want to mention, thank you all for your support, wonderful ideas and especially your smile and laughter, the best medication of all.

Finally, to my beloved family: Father, Abu Bakar Hamzah; Mother, Fatimah Abd. Karim; Brother and Sister In-law, Zaki and Suzana; Sister, Dg. Ratnasari; youngest brother, Iskandar; and especially to Ainol Azifa bte Mohd. Faik, thank you all for your patience, moral and financial support.

ABSTRAK

CHARACTERIZATION AND EXPRESSION OF RECOMBINANT MOUSE GAROUPA (*CROMILEPTES ALTIVELIS*) GROWTH HORMONE & INSULIN-LIKE GROWTH FACTOR GENES

Hormon pertumbuhan (GH) merupakan polipeptida penting yang dihasilkan di dalam kelenjar anterior pituitari yang diperlukan untuk pertumbuhan dan perkembangan normal vertebrata. Penghasilan faktor pertumbuhan insulin I (IGF-I) pula dikawal oleh GH pada peringkat transkripsi di dalam hati. Salah satu ikan marin utama yang dikulturkan adalah kerapu tikus (*Cromileptes altivelis*). Cabaran utama dalam pengkulturan ikan ini adalah kadar tempoh masa tumbesaran yang lebih perlahan berbanding ikan kerapu komersil yang lain. DNA komplementari (cDNA) pre-pro GH dan IGF-I telah dipencilkan menggunakan kaedah amplifikasi hujung cDNA cepat (RACE) terubahsuai, yang merupakan kaedah berdasarkan homolog RACE bersama tindakbalas rantaian polimerase (HRACE-PCR). DNA komplementari GH mengandungi 920 pb di mana 615 pb merupakan jujukan bacaan terbuka (ORF) yang mengkodkan 204 asid amino. Sementara itu, jujukan separa cDNA IGF-I mengandungi 496 pb yang mengkodkan 150 asid amino yang mengandungi domain B, C, A, D dan sebahagian domain E. Untuk tujuan pengekspresan dan pencirian hormon rekombinan ikan kerapu tikus (rmGH), segmen cDNA GH yang mengkodkan asid amino matang telah diselitkan ke dalam vektor pengzahiran *Escherichia coli*, pTrcHis (Invitrogen), dan dikawal oleh promoter *trc*. Induksi isopropyl ß-thiogalactosidase (IPTG) pada kepekatan 1.0 mM dengan suhu 37ºC telah menghasilkan protin bersaiz 23 kDa. Pengesanan protin yang mengandungi histidin mengunakan pemblotan Western di dalam gel menunjukan bahawa protin bersaiz 23 kDa tersebut mempunyai penanda 6X histidin. Hormon pertumbuhan rekombinan ikan kerapu telah diekspreskan sebagai badan inklusi, oleh yang demikian, proses-proses seperti pemencilan, penulenan dan pelipatan semula protin diperkenalkan untuk melipat semula protin. Bagi mengatasi masalah ini, satu proses ekonomikal, boleh diskalakan dan cepat telah dibangunkan menggunakan kaedah dialisis detergent. Penggunaan vektor pTrcGHis didapati menghasilkan 19-21% protin sel *E. coli* sebagai rmGH. Kajian bioasai awalan untuk melihat kesan pemberian rmGH kepada anak ikan tilapia secara oral dua kali seminggu pada kepekatan 1 ng/g, 10 ng/g dan 100 ng/g mengikut berat badan menunjukan bahawa ikan yang menerima dos tertinggi adalah 50% lebih berat ($P<0.05$) dan 10% ($P<0,05$) lebih panjang dibandingkan dengan kawalan, pada minggu ke enam. Keputusan ini menujukkan bahawa rmGH yang disediakan adalah bioaktif dan pelipatan semula rmGH telah dicapai.

ABSTRACT

CHARACTERIZATION AND EXPRESSION OF RECOMBINANT MOUSE GAROUPA (*CROMILEPTES ALTIVELIS*) GROWTH HORMONE & INSULIN-LIKE GROWTH FACTOR GENES

The growth hormone (GH) is an important polypeptide produced in the anterior region of the pituitaty gland which is required for normal growth and development of vertebrates.The insulin-like growth factor 1 (IGF-I), on the other hand is regulated by GH at the transcriptional level in the liver. The mouse garoupa is an important cultured marine finfish. One of the major challenges in rearing this fish is its slow growth rate compared to other commercial grouper species. A pre-pro GH and a partial IGF-1 complementary DNA (cDNA) was isolated using a modified rapid amplification of cDNA ends (RACE) method, which is a homology-base RACE-PCR (HRACE-PCR). The GH cDNA consisted of a 920 bp fragment having a 615 bp open reading frame (ORF) encoding a peptide of 204 amino acids. The partial IGF-1 cDNA consisted of a 496 bp encoding 150 amino acids containing the B, C, A, D and a partial E domain. For expression and production of recombinant mouse garoupa GH (rmGH), a cDNA segment encoding the mature peptide of the pre-GH was inserted into an Escherichia coli expression vector, pTrcHis (Invitrogen) which is controlled by the trc promoter. Induction under 1.0 mM isopropyl ß-thiogalactosidase (IPTG) at 37ºC resulted in the overexpression of a 23 kDa protein corresponding to the predicted size based on amino acid sequence. An in-gel Western blot detection for histidine tagged protein furthermore showed that the 23 kDa band contains a 6X hisitdine tag. The rmGH was expressed as inclusion bodies and was further isolated, purified and refolded. To asses this problem, an economical, scalable and rapid process has been developed based on detergent dialysis. The use of the pTrcGHis construct resulted in the production of an estimated 19-21% of the E. coli cellular protein as rmGH. A preliminary bioassay to study conducted on the effects of oral administration of rmGH to tilapia fingerlings administrated twice a week at concentrations of 1 ng/g, 10 ng/g and 100 ng/g body weight resulted in fishes administered with the highest dosage gaining a 50% ($P<0.05$) increase in weight and 10% ($P<0.05$) in length when compared to the control over a period of six weeks. These results indicated that the rmGH was biologically active and refolded.

TABLE OF CONTENTS

PAGE

PAGE

PAGE

PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

%	percent
°C	degree of Celsius
µg	microgram
µg/µl	microgram per microliter
µg/ml	microgram per milliliter
µl	microliter
aa	amino acid
ACTH	adrenocorticotropin
APS	ammonium persulfate
AVT	arginine vasotocin
BLASTN	basic local alignment search tool for nucleotide
BLASTX	basic local alignment search tool for protein
bp	base pair
BSA	bovine serum albumin
$CaCl_2$	calcium chloride
CBB R-250	Coomassie blue R-250
cDNA	complementary DNA
CLIP	corticotrophin-like intermediate lobe peptide
cm	centimeter
DNA	deoxyribo nucleic acid
dNTP	deoxynuceloside-5'-triphosphate
FSH	follicle stimulating hormone
g	gram
GH	growth hormone
GH-ORF	growth hormone open reading frame
GHRH	growth hormone releasing hormone
GnRH	gonadotropin releasing hormone
GTH-I	gonadotropin hormone I
GTH-II	gonadotropin hormone II
HPA	hypothalamo-pituitary axis
HPI	hypothalamo-pituitary-interrenal axis
H-RACE	homology rapid amplification of cDNA ends
IB	inclusion bodies
IDV	integrated density value
IGFBP	insulin growth factor binding protein
IGF-I	insulin-like growth factor I
IGF-II	insulin-like growth factor II
IGFs	insulin like growth factors
IPTG	Isopropyl-ß-D-thiogalactopyranosid
IRS	insulin receptor substrates
JAK	*Janus* kinase
kb	kilo base
kDa	kilo dalton
kg	kilogram
LB	Luria Bertani
LH	leutenizing hormone
LPL	lipopolysaccaride
mA	mili ampere
m	meter

MAP	mitogen-activated kinase
MBAS	methylene blue active substance
mg	milligram
$MgCl_2$	magnesium chloride
MIP	n-olluscan insulin related peptide
ml	milliliter
M-MuLV	monkey murine leukemia virus
mRNA	messenger ribonucleic acid
MSH	melanotropin
NaOH	sodium hydroxide
ng	nanogram
ng/µl	nanogram per microliter
NPY	neuropeptide Y
$O.D_{600}$	optical density 600
ORF	open reading frame
PCR	polymerase chain reaction
PL	prolactin
POMC	proopiomelanocortin
PRL	prolactin
pTrcGHis	GH constructed in pTrcHis vector
PTTH	prothoracicotrophic hormone
RACE	rapid amplification of cDNA ends
RBS	ribosomal binding site
RM	Ringgit Malaysia
rmGH	recombinant mouse garoupa growth hormone
$rmGHg^{-1}$	recombinant mouse garoupa growth hormone per gram
RNA	ribonucleic acid
rpm	resolution per minute
RT-PCR	reverse transcription – polymerase chain reaction
SDS	sodium dodecyl sulphate
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SH2	*src* homology 2
SL	somatolactin
SRIF	somatotropin release-inhibiting factor
Stat	signal transducers and activators of transcription
STH	somatotropin
Tm	melting temperature
TAE	Tris-acetate-Ethylenediaminetetraacetic acid buffer
TE	Tris- Ethylenediaminetetraacetic acid buffer
TEMED	Tetramethylethylenediamine
TSH	thyroid stimulating hormone
U	unit
UTR	untranslated region
V	voltage

CHAPTER 1

PREFACE

1.1 Introduction

Sabah is one of the eastern states of Malaysia situated at the northern tip of Borneo Island, the third largest island in the world. It is bordered by Sarawak on its southwestern side and Kalimantan (Indonesian Borneo) to the south. Sabah has a coastline of approximately 800 to 900 miles where the South China Sea lies in the west and north, the Sulu Sea in the northeast and the Celebes Sea in the east. Therefore, a huge potential exists for developing the aquaculture industry for the country by exploiting the vast coastal line.

Currently, the demand and value of, live reef fish for food, particularly the groupers (Family Serranidae), have grown markedly in the last two decades, especially in parts of Southeast Asia. In 1997, live fish traded in this region was estimated at about 53,000 metric tons, comprising approximately 30,000 metric tons of grouper (based on figures calculated for Hong Kong and assuming that Hong Kong and China represents about 60% of regional trade) (Johannes and Riepen, 1995; Lau and Parry-Jones, 1999). To meet this demand, two-thirds are met by captured fisheries of market-sized fish; the rest is from 'cultured' fish. These cultured fish are grown from caught smaller wild fingerlings or juveniles.

To ensure future viability of the trade in this family of reef fishes, three main issues have been identified. Firstly, is the vulnerability of the groupers, due to exhaustive fishing and based on records of overexploitation of grouper (Cesar *et al.*, 2000; Sodovy and Vincent, 2001). Secondly, is the destructive methods used in fish practices especially usage of cyanide and bombing which causes a danger to reef associated species that depends on the reef for shelter and food and other reef inhabitants (Johan and Riepen, 1995; Barber and Pratt, 1997). Thirdly, the concern connected to human health; due to the depletion of market-sized fish within Southeast Asia, buyers have searched into the Indian and Pacific Oceans for new seed supplies and, unwittingly at first, brought back to major consumption centers fishes that bear naturally occurring ciguatoxins (Sadovy, 2001). Therefore, urgent actions are needed to develop alternative sources of grouper to lift pressure off wild stocks, to reduce the use of cyanide-sourced fish and to provide safe, ciguatera-free fish.

Aquaculture is one of the major solutions. However, high value marine fish culture such as grouper is still at its infancy where wild seeds are used in most culture. Due to this fact, scientists have been studying ways how biotechnology can increase the production of marine food products, making aquaculture - the growth of aquatic organisms in a controlled environment. This field will allow identification and combination of traits in fish and shellfish to increase productivity and quality. Biotechnology aquaculture research can propel the development of high-quality, economical seafood while simultaneously protect stressed marine populations. Some of the advantages offered through aquaculture biotechnology are reducing half of the cycle time required to raise fishes. In return, demands can be increased without proportionately increasing the use of coastal waters. Food conversion rates are also increased leading to the usage of fewer natural resources in fish production, thus enhancing sustainability. Finally, acceleration of fish growth will increase the

economics of aquaculture by increasing volume of production and produce positive outcomes to land-based systems making them economically viable and competitive when compared to ocean-pen systems.

1.2 Research Background

The growth hormone (GH) is a polypeptide of approximately 22 kDa, synthesized in the anterior region of the pituitary gland. Many tissues have been found to have receptors for GH mainly the liver, where induction of insulin-like growth factor I (IGF-I) occurs (Chen *et al.,* 1994; Shepherd *et al.,* 1997). The main regulatory function on growth-promoting effect has been well-documented, but the physiological role of GH is still being studied (Agellon *et al.,* 1998; Schulte *et al.,* 1989). Effects of GH on osmoregulation has been studied and demonstrated in salmonids (Smith, 1956; Komourdjian *et al.,* 1976; Bolton *et al.,* 1987; Johnson and Björnson, 1994). However, in non-salmonid species, this action remains a subject of discussion (Flik *et al.,* 1993; Borski *et al.,* 1994; Auperin *et al.,* 1995; Guillēn *et al.,* 1996; Shephard *et al.,* 1997).

The influence of GH on fish physiology can be used as a tool to assay its biological activity. Several methods have been used to demonstrate the biological activity of GH: *in vivo* by measuring the body weight increase during weekly supplementation of the recombinant protein (Flik *et al.,* 1993; Johnson and Björnson, 1994), and *in vitro* by the sulphate-uptake method employing cerato-branchial cartilage as substrate and radioreceptor assays (Ash, 1977; Duan and Innui, 1990; Duan and Hirano, 1992; Cheng *et al.,* 1995; Shephard *et al.,* 1997). Another possible method is to indirectly measure IGFs induced by the growth hormone following an assay by Northern blot or RNase protection (Chen *et al.,* 1994, Cheng and Chen, 1995; Shephard *et al.,* 1997).

Most fish species have different growth patterns compared to birds and mammals because they continue to grow even after having reached sexual maturity. Because of this they do not achieve commercial weight and length before sexual maturity. Currently, in the market, fishes that achieve commercial weight and length are sexually matured, and as an example these fishes weight more than 500 g. Through biotechnology, growth of commercial fish species can be stimulated (de la Fuente *et al.,* 1996). However, extensive experimental determination of proper GH requirement to improve growth performance in economical important fish species requires tightly regulated process to achieve faster growth.

The gene coding for this hormone has been studied extensively in several mammalian and piscine species because GH is commercially important in the areas of medicine, animal husbandry, fish farming and animal feed formulations. An analysis of GH complementary DNA (cDNA) sequence of a number of species has shown that the coding regions are highly conserved in mammals, and less conserved in fish. However, 3' ends of the cDNAs are highly conserved in a number of vertebrates including fish (Nicholl *et al.*, 1987; Johansen *et al.,* 1989).

Many efforts have been directed to study the effects of exogenous GH on fish physiology. Through the application of genetic engineering, it has been possible to produce recombinant GH from different aquatic organisms (Sekine *et al.,* 1985; Agellon and Chen, 1986; Sato *et al.,* 1988; Rentier-Delrue *et al.,* 1989; Tsai and Tseng, 1992; Pērez *et al.,* 1994; Cheng *et al.*, 1995; Tsai *et al.,* 1995). Exogenous GH can be administered by diet or injected intraperitoneally (Mclean *et al.,* 1994; Flik *et al.,* 1993; Pērez *et al.*, 1994).

Through advances in transgenic technology, it is possible to demonstrate pleiotropic character of GH in tilapia (de la Fuente *et al.,* 1995; Martinēz *et al.,* 1996; Guillēn *et al.*, 1996; Hernāndez *et al.*, 1997) and other finfish (for review see Mclean and Rahman, 1994; de la Fuente *et al.*, 1996).

In this study, we attempt to isolate the mouse grouper (*Cromileptes altivelis*) growth hormone cDNA followed by expression in *E. coli*. The expressed protein is then isolated and characterized.

1.3 Objectives

The objective of this research can be divided into four parts:

1. Isolate the genes encoding growth hormone (GH) and insulin-like growth hormone factor 1 (IGF-I) from their specific tissues, pituitary and liver.

2. Conducting analysis by comparing the nucleotide sequences of the GH and IGF-I cDNA from *Cromileptes altivelis* with other species recorded.

3. Examine the structural integrity and biological activity of the recombinant peptides. An expression study was conducted by utilizing a prokaryotic expression system and determining the protein size and stability during expression. Since the recombinant mouse garoupa growth hormones (rmGH) were expressed as inclusion bodies, a protein refolding technique was developed to refold the rmGH into its active form.

4. To study the effect of administrating rmGH on mouse grouper physiology by measuring its effect on growth rate (based on weight and length). This was performed by introducing different doses of *E. coli*-derived purified recombinant growth hormone.

1.4 Research Approach

There are two approaches towards achieving the objectives. The first approach is through the design of conserved or degenerate primers coding for the GH or IGF-I from available gene bank sequences in fish. The total RNA isolated will undergo a reverse transcription-polymerase chain reaction (RT-PCR) process to obtain a full-length cDNA. This is the clone and analyzed. This method is also known as homology-based rapid amplification of cDNA ends (H-RACE) developed by Venugopal *et al.* (2000) for cloning of growth hormone encoding cDNA from Indian major carps.

The second approach uses random amplification of cDNA ends and polymerase chain reaction (RACE-PCR). In RACE-PCR, after the cDNA synthesis, conserved sequences are used as forward primers, and the poly T are used as reverse primers generating randomly amplified cDNA fragments. The randomly amplified DNA fragments will be cloned and screened repeatedly by specific probes to specifically obtain the desired DNA which codes the GH or IGF-I.

The amplified DNA from either RT-PCR or H-RACE-PCR will be cloned and sequenced to verify the DNA products and further characterized. The steps will involve expression of the gene of interest in an expression vector and hosts to obtain the expressed protein and then analyzed, under SDS-PAGE, to determine the

molecular weight of the product. Western blot analyses will be done to verify the expressed product.

1.4.1 Random Amplified cDNA Ends (RACE) Strategy

Various cloning strategies have been developed for cDNA (Gubler and Hoffman 1983; Okayama and Berg 1982) with the main objective of obtaining the full-length cDNA copies of scarcely abundant mRNA (Frohman *et al.,* 1988) due to fact of unstable structure and prone to degrade by RNases. Conventional procedures using cDNA library construction required at least 5 mg or more poly A^+ RNA (e.g. 10 mg, Rivolta and Wilcox 1995), or a large amount of total RNA (Schraml *et al.,* 1993). This strategy is confronted with a major task (Ghosh, 1996) when limited amounts of scarce cell lines or tissues, like fish pituitary, are involved. Apart from that, conventional cDNA library construction procedures are cumbersome and involve screening thousand of recombinant phages (Sambrook *et al.,* 1989), which is time consuming and laborious.

Alternatively, Frohman *et al.* (1988) and Ohara *et al.* (1989) have developed a novel and relatively simple polymerase chain reaction (PCR) based method, namely RACE and one sided polymerase reaction. This requires a small amount of total RNA or mRNA, the information of which is copied into DNA by reverse transcriptase, and subsequently amplified by PCR with either gene specific (Frohman *et al.,* 1988) or degenerate (Ohara *et al.,* 1989) primers to generate the required quantities of DNA for cloning into plasmid vectors.

Lemaire *et al.* (1994) had adopted the method of Ohara *et al.* (1989) to successfully amplify GH cDNA from *Pangasionodon giga*s, the giant freshwater catfish, but as two separated DNA fragments. These two DNA fragments must be