The search for cathelicidins in *Gasterosteus aculeatus*

Guðrún Ýr Ásgeirsdóttir

Faculty of Life and Environmental Science
University of Iceland
2010

The search for cathelicidins in *Gasterosteus aculeatus*

Guðrún Ýr Ásgeirsdóttir

12 ECTS credit essay that is a part of *Baccalaureus Scientiarum* degree in Biology

Supervisors
Prof. Guðmundur Hrafn Guðmundsson, Ph.D.
Valerie Helene Maier, Ph.D.

Faculty of Life and Environmental Science
School of Engineering and Natural Science
University of Iceland
Reykjavík, Febuary 2010

The search for cathelicidins in *Gasterosteus aculeatus*

12 ECTS credit essay that is a part of *Baccalaureus Scientiarum* degree in Biology

Faculty of Life and Environmental Science
School of Engineering and Natural Science
University of Iceland
Sturlugata 7
101 Reykjavík

Telephone: 525 4000

Skráningarupplýsingar:
Guðrún Ýr Ásgeirsdóttir, 2010, The search for cathelicidins in *Gasterosteus aculeatus*, BS essay, Faculty of Life and Environmental Science, University of Iceland, 28 pages.

Prentun: Háskólafjölritun / Háskólaprent, Fálkagata 2, 101 Reykjavík
Reykjavík, January 2010

Hér með lýsi ég því yfir að ritgerð þessi er samin af mér og að hún hefur hvorki að hluta né í heild verið lögð fram áður til hærri prófgráðu.

I hereby declare that this thesis is supported by my research work, written by myself and has not partly or as a whole been published before to higher educational degree.

_____

Guðrún Ýr Ásgeirsdóttir

Abstract

Cathelicidin are antimicrobial peptide that have been found in many diverse mammals as well as in fish and chicken. The fact that it can be found in different classes of animals indicates that this gene is important for organism so that it will be selected for. Recent studies have shown that cathelicidin in mammals is at the front line in defending its host to microbe invasions. Its role in fish has not been well established, though studies on cathelicidin in rainbow trout, Atlantic salmon and Atlantic cod strongly indicate that cathelicidin has a role in innate immunity for fish and is well conserved between species.
The aim of this project was to discover a cathelicidin gene in the stickleback species (*G.aculeatus*). The presence of cathelicidin had been indicated by a previous study, and if cathelicidin was found a further aim was to isolate that gene for further studies. This did not prove possible, since a successful detection was not achieved. It is quite likely that the previous study was tainted with salmon cDNA, which would give a false positive for stickleback cathelicidin gene. Also, both reference samples (cod and salmon) were infected with bacteria, but infection of an organism provides increased amounts of cathelicidin. This was not done for the stickleback and therefore might have given a false negative, compared to the cod and salmon signal. In conclusion the methods used in this study, did not prove to be powerful enough to detect and isolate the gene.

Útdráttur

Cathelicidin eru bakteríudrepandi peptíð sem hafa fundist í fjölmörgum mismunandi spendýrategundum ásamt nokkrum fiska tegundum og svo kjúklingum. Sú staðreind að peptíðin finnist í svo mörgum fjarskyldum tegundum dýra gefur til kynna að peptíðið gegni mikilvægu hlutverki fyrir lífveru og því sé það undir mjög sterku, jákvæðu vali í þróun. Nýlegar rannsóknir hafa sýnt fram á að cathelicidin í spendýrum eru mikilvægur hluti af fyrstu vörn þeirra gegn örverum. Þó að hlutverk þeirra í fiskum sé ekki jafn vel skilgreint, þá gefa *in vitro* rannsóknir á bleikju, laxi og þorski sterklega til kynna að þau gegni einnig hlutverki í ónæmiskerfi fiska, og eru mjög vel varðveitt á milli tegunda.
Markmið þessa verkefnis var að staðfesta tilvist cathelicidin gens í hornsýli (*G.aculeatus*), en það hafði áður verið sýnt fram á, með óskyldri rannsókn. Einnig, að ef tækist að sýna fram á tilvist þess, að einangra genið fyrir frekari rannsóknir. Ekki tókst að sýna fram á að cathelicidin genið væri til staðar í hornsýli, þrátt fyrir fyrri niðurstöður. Mjög líklegt er að þessar fyrri niðurstöður hafi verið mengaðar með laxa cDNA sem gæfi falska jákvæða svörun. Einnig voru bæði samanburðarsýnin sem notuð voru í þessu verkefni, lax og þorskur, sýkt með bakteríu en ekki hornsýlis sýnin. Þar sem fjölmargar rannsóknir hafa sýnt fram á að sýking lífveru með bakteríu veldur mikilli aukningu á tjáningu cathelicidin, gæti þessi mismunur á milli hornsýlis sýna og samanburðasýna valdið falskri neikvæðri svörun. Því eru niðurstöður þessa verkefnis að þær aðferðir sem notaðar voru séu ekki nógu sterkar og næmar til þess að staðfesta tilvist cathelicidin í hornsýlum.

Table of Contents

Figures

Tables

Abbreviations

AC	Atlantic cod
AMP	Antimicrobial peptide
AS	Atlantic salmon
BCP	Bromochloropropane
CAMP	Cathelicidin antimicrobial peptide
CRAMP	Cathelicidin-related antimicrobial peptide
cDNA	Complementary DNA
dNTP	Deoxyribonucleotide
EDTA	Ethylenediaminetetraacetic acid
e.g.	exempli gratia
EtBr	Ethidium bromide
EtOHE	thanol
NaCl	Natrium chloride
NaOAc	Sodium Acetate
PCR	Polymerase chain reaction
PMAP-36 and 37	porcine myeloid antimicrobial peptides 36 and 37
PF-1	Prophenin-1
PG-1 to 5	Protegrin-1 to 5
PR-39	Proline-rich antimicrobial peptide
RT	Room temperature
rtCath1	Rainbow trout cathelicidin 1
rtCath2	Rainbow trout cathelicidin 2
SB	Stickleback
SDS	Sodium dodecyl sulfate

Acknowledgment

I would like to thank Professor Guðmundur Hrafn Guðmundsson for giving me the opportunity to work on this project. I would also like to thank Valerie Helene Maier for all her guidence and patience throughout this project, always by my side answearing questions and willing to help in any way possible. Finally I would like to thank my family for their support and understanding these past few months.

1 Introduction

1.1 Antimicrobial peptides

Antimicrobial peptides (AMPs) have been shown to display a broad spectrum of antimicrobial activity and to have a role in the innate immunity system, such as to provide a first defence against microbes. When the amino acids composition, size and conformational structures of various antimicrobial peptides are compared, it is possible to categorize them into four general classes (Figure 1).

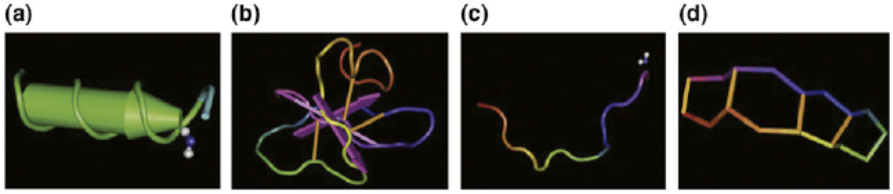

Figure 1 Four general structural classes of AMPs. a) structure of LL-37 core peptide (cathelicidin) bound to detergent micelles, b) structure of hBD2 (defensin), c) structure of bovine antimicrobial peptide indolicin bound to DPC micelles, and d) structure of cyclic defensin from a rhesus macaques leukocytes (Lai and Gallo, 2009).

AMPs can therefore have the structure of an α-helix (fig.1.a) which is a very common form of these peptides, such as Bombinin H4 peptide from amphibians; a β-sheet stabilized with cysteine bridges (fig.1.b) such as BTD-1 isolated from a Baboon; an extended structure (fig.1.c) or a loop structure (fig. 1.d). (Lai and Gallo, 2009).

A major part of antimicrobial peptides seem to be relatively small in size (12-15 amino acids), are cationic and have an amphipathic conformation (Lai and Gallo, 2009). These characteristics are fundamental to the function of the peptides. They target the membrane or envelope of microbes, but do not attack membranes of multicellular organisms (Figure 2).

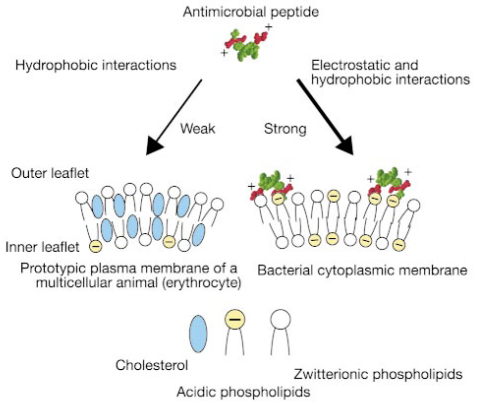

Figure 2 A simple schematic drawing of how AMPs distinguish between self and non-self membranes. The positively charged peptide is attracted to the negatively charged lipid heads on microbe membranes, whereas multicellular membranes have a neutral charge on the outside, and their negatively charged lipids buried on its inner side, where the peptide can not access them (Zasloff, 2002).

This discrimination between self and non-self is acchieved by the peptides using the fundemental difference between the design of microbe and multicellular animal membranes. This is best understood and explained with relation to bacterial membranes, which have their outermost area of the bilayer populated with negatively charged lipids, whereas multicellular membranes usually have un-charged lipids in their outermost area. The Shai-Matsuzaki-Huang (Zasloff, 2002) model explains the activity of most antimicrobial peptides. It proposes that the peptide interacts with the membrane by attracting to the negative charge, which causes the displacement of lipids and therefor the alteration of membrane structure and in some cases the entry of the peptide into the cell. There have been many hypotheses proposed about how these AMPs actually kill microbes. These hypothesis include fatal depolarization of the bacterial membrane and the creation of holes that cause cellular contents to leak out, among others (Zasloff, 2002).

These peptides have been demonstrated to kill Gram-negative and Gram-positive bacteria, mycobacteria, enveloped viruses, fungi and even cancer cells (Tossi, 2005). More than 800 types of peptides have been characterized from diverse animal classes, ranging from mammals to fish to chicken. These peptides have mainly been found in two locations in mammals, on epithelial surfaces were they serve as a local defense mechanism, and in circulating myeloid derived cells which gather at the site of microbial invasion. They are dividied into several families, such as cecropin (in insects), magainins (in amphibians), defensins and cathelicidins (in mammals, bird and fish) (Linde et al., 2009).

1.2 Cathelicidin

A specific antimicrobial peptide family that has been widely studied is the cathelicidin family. All known cathelicidins are divided into three structural classes (Figure 3) very similar to those structural classes antimicrobial peptides are categorised into (as discussed above).

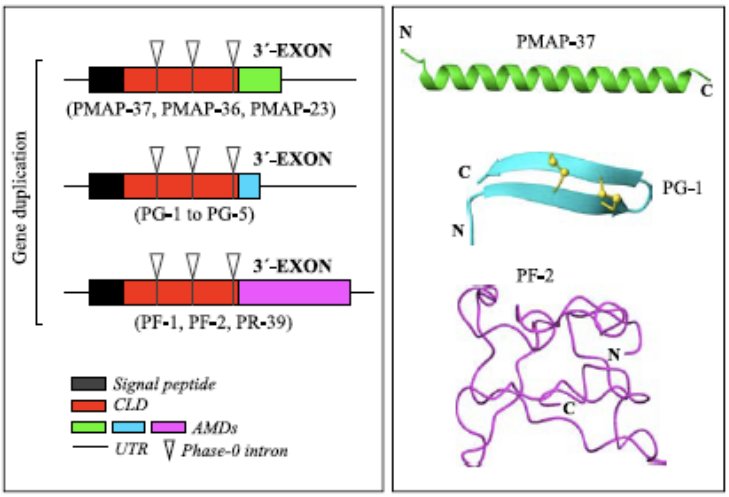

Figure 3 The structural diversity of AMPs from paralogous cathelicdins. Left: gene structure with precursor signal peptide, cathelin-like domain (CLD) and the peptide (AMD). Right: Structure of different antibacterial peptide classes (Zhu and Gao, 2009).

These three classes were the process of many gene duplication events, and as demonstrated in the figure above, are linear-helical peptides (e.g. PMAP-37, PMAP-36), peptides with disulfide bridges and β-sheet structures (e.g. PG-1 to PG-5) and peptides enriched for

specific amino acids (e.g. PF-1 and PR-39). This diversification can provide the host with a multidrug and synergistic defense system, but it is not known how the different peptides originated after gene duplication event (Zhu and Gao, 2009).

All the genes that belong to the cathelicidin family have been shown to have a highly conserved preproregion that contains the cathelin-like domain, which has a protein from the N-terminus and a small, variable antimicrobial peptide at the C-terminus. This preproregion can be further divided into four exons, where exon one encodes the signal peptide, exon two and three encode the cathelin region, and exon 4 mainly encodes the active peptide region. The variable peptides is released by a enzyme-mediated cleavage, and like the majority of antimicrobial peptides, will in most cases have an amphipathic structure and a positive charge. Another similarity among different cathelicidin is the cysteine motif, but almost all cathelicidin have been shown to have four cysteines, which in mammals are within the motif C-X_{10}-C-X_5-G-X_4-C-X_{16}-C, clustered in the cathelin like domain (Chang et al., 2005).

Cathelicidin do not only take part in the innate immunity system by microbicidal activity and neutrolizing endotoxins, but also to recruit neutrophils, monocytes and T cells to sites of microbe infection through a FPRL1 receptor. They therefore link the innate immunity system to the adaptive immunity system (Lai and Gallo, 2009). The number of cathelicidins present in a species varies, in pigs (*Sus scrofa*) they are at least 12, 7 in cattle (*Bos taurus*) and similar for sheep. Only two species have been shown to contain only one cathelicidin gene, which is the CAMP gene in humans (*Homo sapiens)* and the CRAMP gene in mice (*Mus musculus*) (Zhu, 2008).

Not that long ago it was thought that mammals were the only class of animals to have cathelicidin as a part of their innate defense mechanism. Then in 2003 Uzzell and colleagues published their discovery of two novel sequences in Atlantic hagfish (*Myxine glutinosa*) that are ancient members of the cathelicidin gene family. This indicated that antimicrobial peptides are an ancient class of peptides that originated before separation of fish and vertebrates (Uzzell et al., 2003).

1.3 Cathelicidins in fish

Fish is the earliest class of vertebrates that has both innate and acquired immunity, even though their adaptive immunity is not as well developed as that of higher vertebrates. They are in constant contact with various types of microbes and have to be able to defend themself, which explains the high amount of antimicrobial peptides found in their epithelial cells and mucosal surfaces. Most species of fish have many other innate defense molecules at those sites to, such as apolipoprotein and lysozyme to name a few (Plouffe et al., 2005). This innate protection is non-specific, fast and relatively temperature independent. This is very important for marine animals, since the adaptive immune system can take some time to respond, and is temperature dependent (Ellis, 2001).

Many different types of antimicrobial peptides have been discovered in various fish species, such as paradaxin (found in Moses sole fish), pleurocidin (in flounder), misgurin, piscidin and then cathelicidin (Plouffe et al., 2005). When these fish cathelicidin are considered carefully, it becomes apparent that they have some variations from the general cathelicidin structure in vertebrates, such as that the signal peptide is shorter, while the

cathelin site is longer. The cysteine motif though is very similar to the vertebrate motif, or C-X_{10}-C-X_5-G-X_4-C-X_{23}-C (Chang et al., 2005).

Chang and colleagues discovered that a rainbow trout (*Oncorhynchus mykiss*) has two cathelicidin genes, and they showed that a synthesized predicted fragment of rtCath_1 displayed in vitro antimicrobial activity. RtCath_1 was only expressed in detectable amount in gill, head kidney and spleen after infection and rtCath_2 was always expressed in gill, head kidney, intestines, skin and spleen but upregulated in these tissues after infection. They also showed the existence of two cathelicidin genes in Atlantic salmon (*Salmo salar*) which corresponded to the two genes found in rainbow trout (Chang et al., 2005 and Chang et al., 2006). More cathelicidin genes were discovered in salmonidae species soon after that (Maier et al., 2008) where they followed the same general model. The only divergence was that exon three in Arctic char (*Salvelinus alpinus*) and in brook trout (*Salvelinus fontinalis*) of one of the genes was missing, which indicates a single deletion event when the *Salvelinus* family diverged from the other salmonidae families. Maier and colleagues also discovered that Atlantic cod (*Gadus morhua*) had three cathelicidin genes, which follow the general motif for fish cathelicidin genes, but other than that, diverge from other discovered fish cathelicidins. They also showed that cod cathelicidins are all very cationic which suggests a strong antibacterial function of the peptide (Maier et al., 2008).

1.4 Three-spined stickleback (G.aculeatus)

The three-spined stickleback species is a phenotypically diverse population, where even some populations have reached the mark where they are characterized as a biological different species. This makes the species optimal for analysis of evolutionary mechanisms, and since it has a highly ritualized behaviour it has already been adapted as a model species for behavioural ecology (Bell and Foster, 1994).

The species is considered to have many attributes that make it a good model species for ecology and evolutionary biology, since they are easy to catch from the wild and breed in a lab. Their phenotypic differences make them ideal for comparative studies, they also produce many offspring and they have a short generation time so they can adjust fast to a new environment. Since they are vertebrates, it might even be possible to use them as a model species in genetics for all vertebrates (Mbl.is, 2004).

Many studies have shown that the three-spine stickleback species is plagued by parasites, and some researchers have even suggested that these parasite infections play an important role in the species divergence, saying that when the sticklebacks adapt to a new environment (e.g. parasites) they may form a mating barrier between ecotypes (Scharsack et al., 2007).

1.4.1 Phylogeny of stickleback, salmon and cod

When the phylogeny of salmon (*S.salar*), cod (*G.morhua*) and stickleback (*G.aculeatus*) are compared to each other, with the data available today, it seems that stickleback is more closely related to cod, than salmon or other salmonids. The last known, joined class of salmon and stickleback is Euteleostei that diverged in the Triassic period (251-199 Mya). On the other hand the last known joint class of cod and stickleback appears to be Acanthomorpha which is a subgroup of Euteleostei (The Tree of Life web project).

1.5 The aim of the project

The aim of this project was to verify results from an earlier research, which indicated that salmon cathelicidin primer pairs bound to a cathelicidin gene in *G.aculeatus* (data not shown). A further aim was to discover and isolate a cathelicidin gene in *G.aculeatus* with the aid of degenerate primers that were constructed from known cathelicidin sequences of other fish species (see Appendix B).

2 Materials and methods

For more detailed description of methods, see Appendix A.

2.1 Sampling

2.1.1 Atlantic salmon (Salmo salar)

Atlantic salmon was previously infected with *Moritella viscosa* and the researcher was provided with cDNA, which was acquired with methods described in sections 1.2 and 1.4.

2.1.2 Atlantic cod (Gadus morhua)

Atlantic cod was infected with *Aeromonas salmonicida* previously and as before, researcher was provided with cDNA that had been acquired with methods described in sections 1.2 and 1.4

2.1.3 Three-spined stickleback (Gasterosteus aculeatus)

The sticklebacks used in this study were collected at different times from different locations around Iceland.

Sticklebacks caught in Héðinsfjarðarvatn in Héðinsfjörður 2008 were stored whole in ethanol at room temperature for DNA extraction, and stored in RNAlater (Ambian, Austin, TX, USA) at -20°C for RNA work.

Stickleback caught from the river in Vatnsmýrin 2009, was frozen whole in liquid nitrogen and stored at -80°C

Stickleback caught in Fljótshlíð 2009 was dissected and the internal organs were frozen in liquid nitrogen and stored at -80°C.

2.2 RNA extraction

Stickleback samples were homogenized in TRI Reagent® (Ambion, Austin, TX, USA) with a PELLET PESTLE® Cordless Motor (Kimble-Kontes, Vineland, NJ, USA) and RNA extracted according to manufacturers protocol. Genomic DNA was digested with DNase I (New England Biolabs, Ipswich, MA, USA), followed by ethanol precipitation of the RNA.

2.3 DNA extraction

Genomic DNA was isolated from stickleback, cod intestines and salmon intestines with a modified phenol-chloroform DNA extraction protocol using a lysis buffer with the final molar concentration of 10 mM Tris, 0.1 M EDTA and 0.5% SDS (Appendix A). For the

salmon tissue the incubation was continued over night with an extra 500 µl of lysis buffer and 2.5 µl of proteinase K.

2.4 Nucleic acid quality control

Quantity and quality of both RNA and DNA was evaluated using NanoDrop® ND-1000 UV-Vis Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). The quantity of the nucleic acids (ng/µl) was calculated as 40x the absorption at 260 nm. The quality of the nucleic acids was measured with A260/A230 and A260/A280 ratios. The former should be around 1.8 for pure DNA and 2.0 for pure RNA, the latter should be approximately 2.0-2.2 for both. Lower values give an indication that there is contamination in the sample.

2.5 cDNA synthesis

1000 ng of RNA were reverse transcribed into cDNA using First Strand cDNA synthesis kit (Fermentas #1612) with random primers according to manufacturer's protocol.

2.6 PCR

For the PCR reactions, different primer sets (Table 1, 2 and 3) with various templates were used and a basic PCR protocol applied with slight variations in annealing temperature, as indicated by X.

When salmon primers were used on stickleback cDNA the annealing temperature (X) was 51°C. When degenerate primers were used, gradient PCR between 50-60°C was performed to determine the optimal temperature for the primers. The gradient PCR indicated that the optimal annealing temperature for all of the degenerate primers was close to 54°C, which was then used.

2.6.1 Gel electrophoresis on PCR products

PCR products were separated on a 1,2 % agarose gels using TAE buffer and visualised with ethidium bromide on a UV light board.

2.6.2 Expected band size for various primer pairs

To increase the probability that the bands visualized on an agarose gel were for the cathelicidin sequence, expected sizes for all three types of primers were calculated.

Table 2-1 Expected band size for Atlantic salmon primers

	R1	R2	R3
F1	178	322	383
F2	142	286	347
F3	-	46	107

Table 2-2 Expected band size for Atlantic cod primers

	codR1	codR2	codR3
codF1	233	369	596
codF2	197	333	547
codF3	71	206	420

Table 2-3 Expected band size for degenerate primers

	fDegR0.5	fDegR1
fDegF1	249-267	357-405
fDegF1.1	152-165	267-303
fDegF1.2	147-159	261-297
fDegF1.3	90-102	224-239
fDegF2	NA	138-159
fDegF3	NA	NA

Primers with a b in the name (not shown here) had the same expected size as primers without.

2.6.3 Gel extraction

Bands were cut out of agarose gel and isolated from the gel with Gen Elute™ Gel extraction kit (Sigma-Aldrich, Taufkirchen, Germany) according to manufacturers protocol. PCR products were eluted from the column with 30 µl of elution buffer.

2.7 Cloning

Cloning was performed using the Invitrogen pCR8 protocol, with the exception of reducing the amount of salt, water and vector for each reaction to half of what is mentioned in the protocol. E.coli cells were then transformed, spread on LB agar plates containing spectinomycin and grown over night at 37°C.

2.8 Plasmid isolation and sequencing

Colonies from the over night plates were picked with a sterile tip and cultured in a LB medium containing spectinomycin, at the same time a PCR reaction was performed with corresponding colonies. The PCR was performed with plasmid primers GW1 and GW2 and annealing temperature of 60°C. The reaction was then electrophorised on an agarose gel and analyzed using a UV light and a camera.

For the plasmid isolation, Plasmid DNA Purification (Nucleospin plasmid) kit (Macherey-Nagel, Düren, Germany) was used on the over night cultures according to the protocol from the manufacturer with the only exception of using only 30 µl of AE buffer to elute the plasmid DNA from the column. The isolated plasmids were then cut with EcoRI restriction enzyme, visualised on a gel and compared with uncut plasmids to determine if the inserted piece was of the right size.

To sequence the cloned pieces a sequencing PCR reaction was performed using BigDye Terminator v3.1 (Applied Biosystems) and GW1/GW2 plasmid primers. The DNA was then precipitated, dissolved in HiDi and analyzed in a ABI sequencing machine. The obtained sequences were analyzed by blasting them on the NCBI webpage (http://www.ncbi.nlm.nih.gov/).

3 Results

3.1 DNA extraction

The quality of the extracted DNA was measured with NanoDrop ND-1000 spectrophotometer which gave high values for all samples. The quality and amount of DNA in samples was confirmed by gel-electrophoresis on a 0.8% agarose gel (Figure 4).

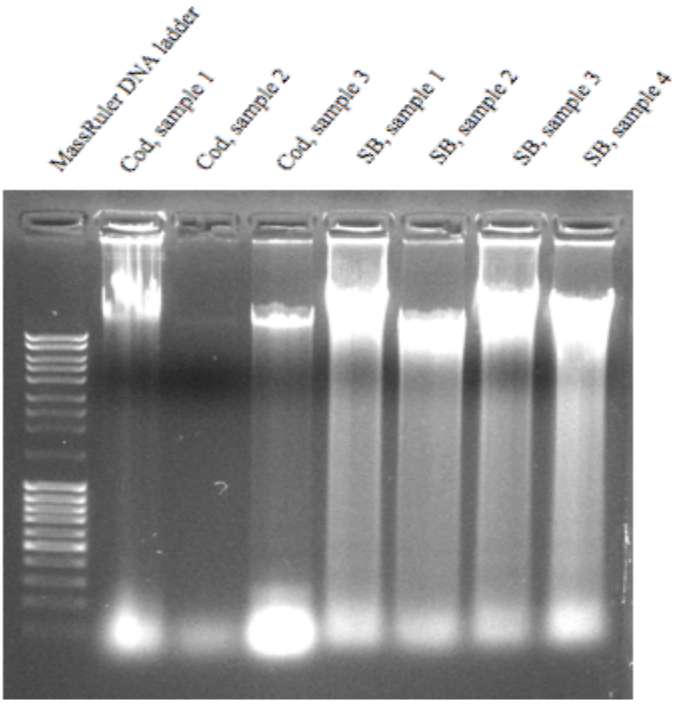

Figure 4 Quality of extracted DNA from AC and SB tissues.

DNA amount and integrity seemed to be good in most samples, although the first cod sample was subjected to shaking during the extraction procedure, which may have resulted in low quality DNA, even though absorption ratios were within normal limits. Cod sample 2 also had lower amount of DNA but was still within usable amounts.

3.2 PCR with salmon and cod cathelicidin primers

As noted before, results from a previous experiment of another researcher indicated that cathelicidin salmon primers bind to stickleback cDNA and when the samples were analyzed with gel-electrophoresis and southernblotting, the bands appeared to be of the correct size according to expected values (Table 1). The PCR was repeated to see if the results were reliable enough for isolation of the gene (Figure 5).

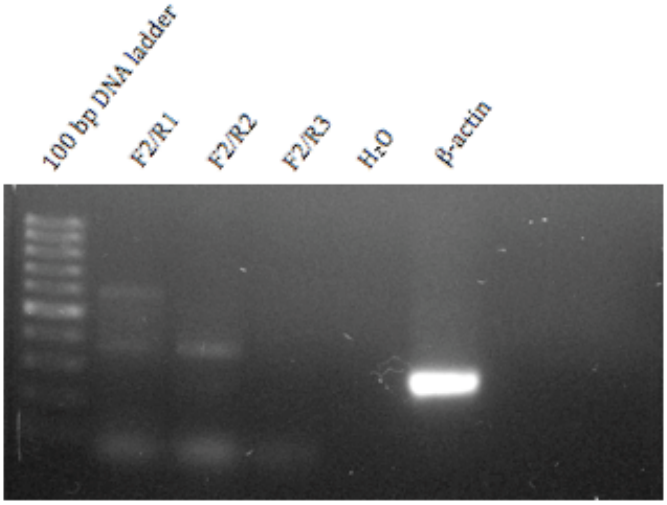

Figure 5 PCR performed on SB cDNA with AS primers F2, R1, R2 and R3. Faint bands are visible for F2/R1 and F2/R2 primers, both seem to be 3-400 bp in size. Water was used as a negative control to detect contamination of samples, and β-actin was used to determine the quality of cDNA. Annealing temperature (X) was 51°C.

When the experiment was repeated, no bands of the correct size were visualized for any of the salmon primer pairs (Figure 4), not even F2/R1 or F2/R2, which seemed to be the most promising once in the earlier experiment. Therefore further uses of salmon cathelicidin primers were abandoned in the search for cathelicidin gene in stickleback, and other options explored. Previously a band of expected size for the salmon primer pair F2/R2 had been extracted from an agarose gel. This band was cloned and sequenced as described below.

Stickleback is distantly related to both the salmonid fish family and cod family, both of whom have the cathelicidin gene defined. Therefore we wanted to determine if cod cathelicidin primers bound to stickleback cDNA at the correct location, when salmon cathelicidin primers did not (Figure 6).

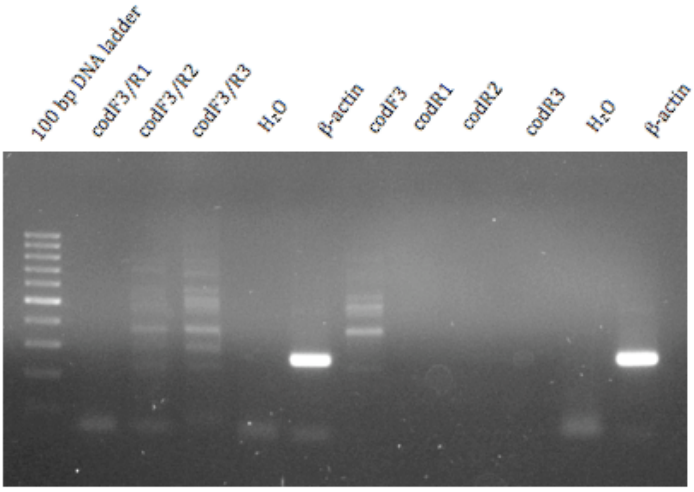

Figure 6 PCR performed on SB cDNA with AC primers. Bands are visible for primer pairs codF3/R2 and codF3/R3 that appear to be of the right size, but those bands are also visible when codF3 primer is used on its own. Water is used as a negative control and β-actin as a positive control. Annealing temperature (X) was 55°C

When the data was analyzed, it became apparent that the codF3 primer on its own produced several bands, some who corresponded exactly to the bands of expected size (Figure 5), and therefore it could not be used for further studies of cathelicidin in stickleback.

3.3 Degenerate primer pairs fDegF1/R1 and fDegF2/R1

Since neither salmon nor cod cathelicidin primers showed promising results for stickleback, degenerate primers were constructed (as described in Appendix B).

3.3.1 PCR with degenerate primer pairs fDegF1/R1 and fDegF2/R1

Primer pair fDegF1/R1 generated two bands on an agarose gel (Figure 7), and primer pair fDegF2/R1 generated weak bands (data not shown).

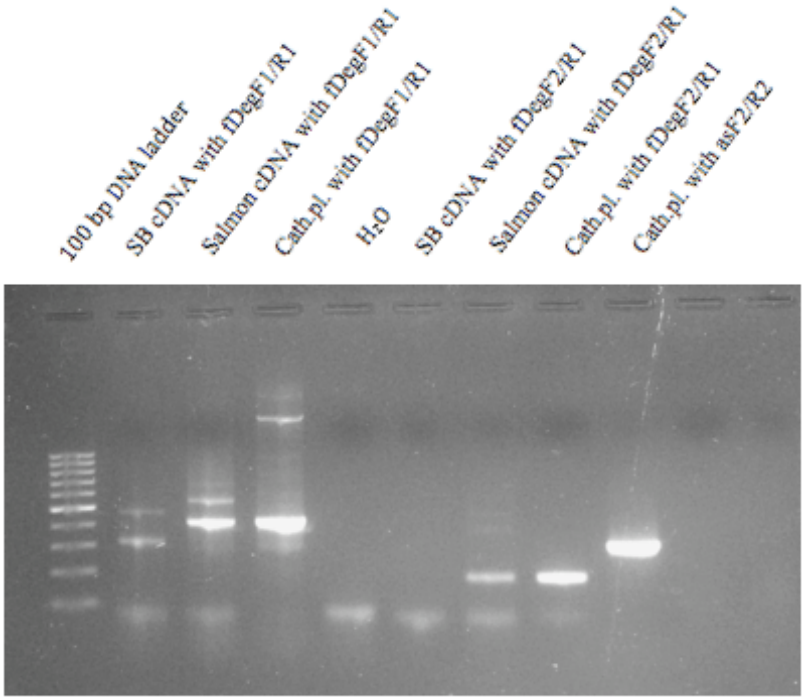

Figure 7 PCR on SB and AS cDNA with degenerate primers fDegF1, fDegF2 and fDegR1. A band corresponding to expected size is seen for primer pair fDegF1/R1, but none for primer pair fDegF2/R1. Water is used as a negative control and β-actin as a positive control. Annealing temperature (X) was 53°C

The size of the lower band for fDegF1/R1 was observed to correspond to expected size (Table 3) and therefore required a closer look. For primer pair fDegF2/R1 three faint bands were visualized, the lowest one having a size of ~150-160 bp, which corresponds well with expected size (Table 3). This band was excised and extracted out of the gel.

Quality of degenerate primers

As a control, a PCR reaction was performed with the degenerate primers, but only one primer was used for each reaction (Figure 8)

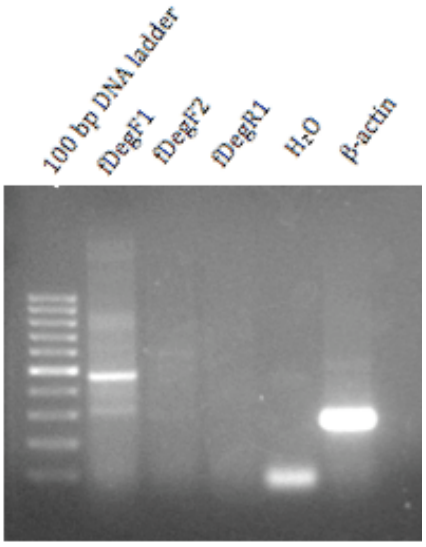

Figure 8 PCR on SB cDNA with only one primer per reaction, fDegF1, fDegF2 and fDegR1. fDegF1 primer shows bands on it's own, but fDegF2 and fDegR1 do not. Water was used as a negative control, and β-actin as a positive control. Annealing temperature (X) was 54°C.

When the data was analysed, it became apparent that primer fDegF1 generated both bands on its own, and therefore was not continued with. On the other hand, neither primer fDegF2 nor fDegR1 gave bands on its own, so it was possible to continue working with bands generated by primer pair fDegF2/fDegR1.

3.3.2 Gel extraction and cloning

The lower band for fDegF2/R1 was purified out of a gel and cloned into a vector. Another band, extracted from a previous experiment with salmon primers F2/R2 was also cloned into vectors (as discussed in section 4.2). These vectors were then inserted into competent E.coli cells and plated onto a LB agar plate containing spectinomycin.

3.3.3 Colony PCR, plasmid isolation and sequencing

After colonies had grown over night, colony PCR was performed to estimate the size of the pieces that were incorporated into the vectors (Figure 9).

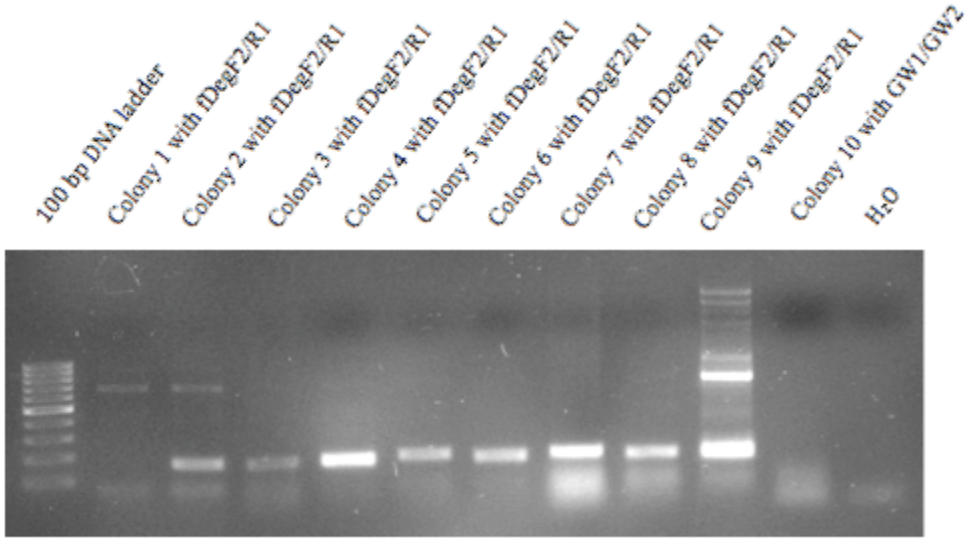

Figure 9 Colony PCR with fDegF2/R1 and GW1/GW2 primer pairs. For colonies 2-9 there are clear bands visible but their size is variable. Water was used as a negative control. Annealing temperature (X) was 54°C.

Plasmids were then isolated from the E.coli cells for sequencing. The plasmids were cut with EcoRI (NEB) for better analysis of the size of inserted pieces (Figure 10)

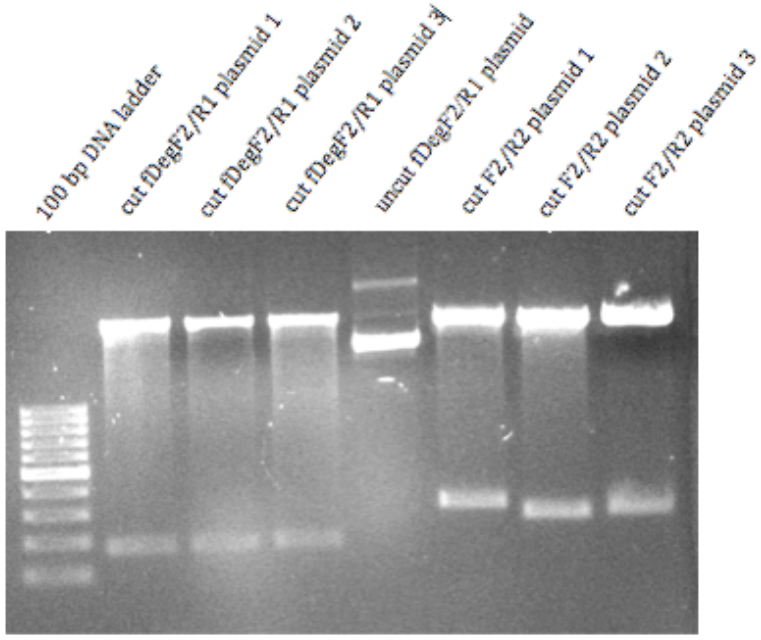

Figure 10 Size estimation of incorporated bands. The bands in plasmid 1, 2 and 3 for fDegF2/R1 all seem to be of similar size. However, the band for cut F2/R2 plasmid 1 seems to be a few base pairs larger than the once for plasmid 2 and 3. The uncut plasmid is used as a comparison.

The higher bands are cut plasmids, while the lower bands correspond to the incorporated PCR fragments. These fragments for primer pair fDegF2/R1 are all of the same size, around 180 bp, although the expected size was around 160 bp. The inserted fragment for F2/R2 plasmid 1, is slightly larger than for plasmid 2 and 3.

A sequencing reaction was performed on both the fDegF2/R1 degenerate primer pair and the F2/R2 salmon primer pair fragment, the results analyzed with ABI sequencer and the

resulting sequences then blasted to the NCBI database. This did not acquire expected results; only one sample got a hit for anything other than unidentified sequences in various organisms. This was plasmid 3 from F2/R2 with GW1 primer where a translated reading frame (nr.1) gave a perfect match to a *Danio rerio* protein. This protein, called dkey-40c11.1, is a homolog of a gene family in mammals, called rhotekin and is important for cellular regulation.

Since these primer pairs did not give any realistic results for cathelicidin gene, these primers were abandoned and the search for a better degenerate primer pair continued.

3.4 PCR with other degenerate primers

More degenerate primers were constructed as before (Appendix B) but these relied more on the nucleotide sequence of cathelicidin than amino acid sequence. A PCR reaction was performed with these primers on SB, AS cDNA and a plasmid containing salmon cathelicidin (Figure 11).

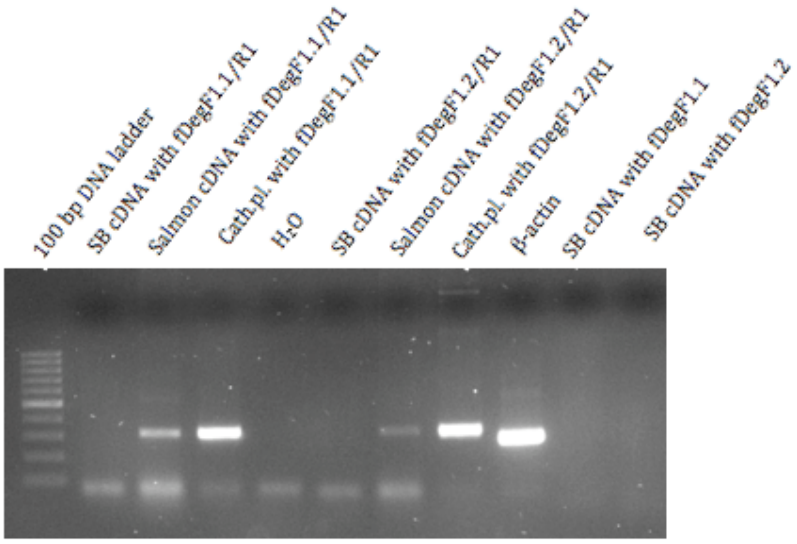

Figure 11 PCR on SB cDNA with degenerate primers fDegF1.1, fDegF1.2 and fDegR1. There seem to be no visible bands for stickleback cDNA with either primer pair but both bind to AS cDNA. Water is used as a negative control, and β-actin as a positive control. Annealing temperature (X) was 54°C.

Neither primer pair fDegF1.1/R1 nor fDegF1.2/R1 bound to stickleback cDNA. These primers bound to salmon cDNA and cathelicidin plasmid and gave fragments of size corresponding with expected size (Table 1), which indicated that the primers bound to the cathelicidin template. The new primers were also tested on their own to see if they amplify fragments without the corresponding reverse primer, but neither did.

The primers were further optimised to bind both salmon and cod cDNA. This was done since the original primers had been constructed almost entirely against salmon. These primers were then used in a PCR reaction on salmon, cod and stickleback cDNA (Figure 12)

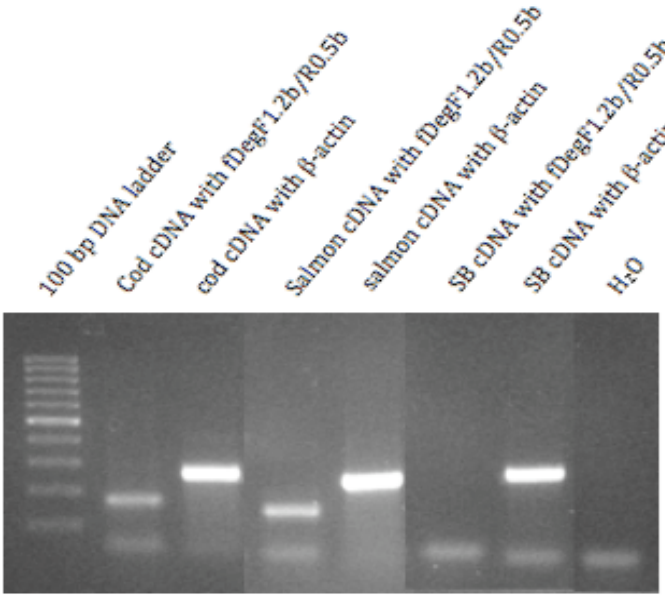

Figure 12 PCR on SB, AC and AS cDNA with degenerate primer pair fDegF1.2b and fDegR0.5b. Primer pair binds both to AC cDNA and AS cDNA, but not to SB cDNA. Water was used as a negative control, and β-actin for each cDNA as a positive control. Annealing temperature (X) was 53°C.

The amplification with primer pair fDeg1.2b and fDegR0.5b produced a band of appropriate size in salmon cDNA and cod cDNA. These primers on the other hand did not bind to stickleback cDNA, which might be due to the fact that these three species are quite distantly related.

3.5 Hidden Markov Model (HMM)

Another approach to finding a cathelicidin gene in stickleback was used by designing a so called Hidden Markov Model. The model was constructed by comparing amino acid sequences of many mammalian and fish cathelicidin for amino acid function (if they are acidic, basic etc.) at similar positions in the sequences, in stead of amino acid type. A factor in the model were also the exon-exon boundaries for every sequence, and if these boundaries were the same among different sequences. When the model was used to scan the stickleback DNA library, it yielded very low probability of the cathelicidin to be present in the library. On the other hand, the model was tested against one cod cathelicidin sequence, and as expected it gave a high probability of finding the cathelicidin gene there.

4 Summary and discussion

Neither salmon nor cod primer pairs were efficient enough to detect the cathelicidin gene in stickleback, even though previous experiments had resulted in an amplification and extraction of a band of expected size. This could be due to contamination of the samples with salmon cDNA or cathelicidin plasmid.

Many degenerate primers were constructed, but only a few primer pairs showed any results. fDegF2/R1 primer pair generated one band that was of expected size which was extracted from the gel. That, and a salmon primer pair band (F2/R2) from a previous experiment were cloned into a vector, their plasmids isolated and the fragments sequenced. When a colony PCR was performed, it became apparent that the inserted fragments were of different size, which was further confirmed by digestion of plasmids with EcoRI endonuclease. This indicated that the inserted fragments could be something other than cathelicidin, since they differed in expected size from fragment actual size. It would have been good to perform a nested PCR on the fragments, where a third primer would have been used to increase the specificity for cathelicidin, since it is highly unlikely that other sequences amplified would bind to a third primer designed for cathelicidin. Unfortunately it was not possible to perform this, since there was not enough conservation between the cathelicidin sequences to construct additional primers between existing primers.

These fragments (12 fragments for fDegF2/R1 and 8 fragments for F2/R2) were then sequenced and blasted against NCBI database, only one sequence got a result for something other than a cDNA library clone for different species. This was plasmid 3 from F2/R2 with GW1 primer and a translated reading frame, which gave a perfect match to a cell regulation protein found in *D.rerio*. If the sequenced fragments (the once that were shown to be quite similar, or even identical to sequenced cDNA clones) were in fact cathelicidin, they should have shown to have the four conserved cysteine residue in the cathelin pro-region (see introduction) similar to already discovered cathelicidin sequences in related species. Therefore, the sequencing of the products from these two primer pairs did not give the expected results.

Other primers were constructed and optimized for binding of both salmon and cod cDNA, since the degeneracy of those primers increased the odds for them to anneal to cathelicidin on SB as well. Primer pair fDeg1.2b/fDegR0.5b produced bands of expected size in both cod and salmon cDNA, but no band for SB cDNA. Therefore it is possible to conclude that this primer pairs do not manage to amplify the cathelicidin in stickleback, even though it seems to amplify cathelicidin for both cod and salmon.

The Hidden Markov Model designed for detecting cathelicidin, yielded a very low probability for the cathelicidin gene being present in the stickleback species. To try the accuracy of the model, it was run on a already sequenced cod cathelicidin and this gave high probability for finding the cathelicidin there. If that had given low probability for finding cathelicidin in that sequence, it would have shown that the model was highly inaccurate and therefore not usable. The strength of the model must be further tested before the results from it can be thought of as reliable. Future studies to test the model could

include searching for cathelicidin in cat by running the model on the cat chromosome DNA which is known to contain the cathelicidin gene. This would prove the model to be strong enough to detect cathelicidin in other species. A second test would include to make a model for lysozyme, which is a gene that has been showed to be very conserved between species and has been discovered in many fish species such as salmon, cod and zebrafish. This would prove the method to be relevant and to be able to find genes in the stickleback genome. If the model will be shown to work, it will not only further sustain the idea that cathelicidin is not present in stickleback, but also makes it possible to search for it in other fish genomes that have been sequenced, such as zebrafish (*D.rerio*) or pufferfish (*Takifugu* species).

All of these results together could indicate that the cathelicidin gene may not be present in the stickleback species studied. It is far more likely though, that the methods used are not powerful enough to detect the gene without more data to rely on. Also, since the sticklebacks sampled were wild, it was not known if the samples were infected. The cod and the salmon sample were infected with different bacteria, and the production of cathelicidin transcription is increased when an organism is infected (Chang et al., 2005; Chang et al., 2006 and Maier et al., 2008). Therefore it would be a recommended to infect a stickleback with bacteria, and then repeat the PCR reactions for salmon, cod and fDegF1.2b/fDegR0.5b primer pairs. If the results are still negative, it may be interpreted as another indication of the absence of cathelicidin in the three-spined stickleback species.

References

Bell, M.A. and Foster, S.A. 1994. *The Evolutionary biology of the threespined stickleback.* Chapter 1. Oxford University Press, USA

Chang, C.I., Pleguezuelos, O., Zhang, Y.A., Zou, J. and Secombes, C.J. 2005. Identification of a Novel Cathelicidin Gene in the Rainbow Trout, *Oncorhynchus mykiss*. *Infection and Immunity* 73(8): 5053-64

Chang, C.I., Zhang, Y.A., Zou, J., Nie, P and Secombes, C.J. 2006. Two Cathelicidin Genes Are Present in both Rainbow Trout (*Oncorhynchus mykiss*) and Atlantic Salmon (*Salmo salar*). *Antimicrobial Agents and Chemotherapy* 50(1): 185-95

Ellis, A.E. 2001. Innate host defence mechanisms of fish against viruses and bacteria. *Developmental and Comparative Immunology* 25(8-9): 827-39

Linde, A., Wachter, B., Höner, O.P., Dib, L., Ross, C., Tamayo, A.R., Blecha, F. and Melgarejo, T. 2009. Natural History of Innate Host Defense Peptides. *Probiotics and Antimicrobial Proteins* 1:97-112

Maier, V.H., Dorn, K.V., Gudmundsdottir, B.K. and Gudmundsson, G.H. 2008. Characterisation of cathelicidin gene family members in divergent fish species. *Molecular Immunology* 45: 3723-30

Mbl.is. 2004. Rannsókn á erfðaeiginleikum hornsíla fer fram á Íslandi: Gefa vísbendingar um gen sem stýra vexti útlima mannsins. Mbl.is/gagnasafn. Website: http://www.mbl.is//mm/gagnasafn/grein.html?radnr=778832;minningar=0;gsmession=7029182192;firstgsm=4148a889. © mbl.is/Árvakur hf.

Plouffe, D.A., Hanington, P.C., Walsh, J.G., Wilson, E.C. and Belosevic, M. 2005. Comparison of select innate immune mechanisms of fish and mammals. *Xenotransplantation* 12(4): 266-77

Scharsack, J.P., Kalbe, M., Harrod, C. and Rauch, G. 2007. Habitat-specific adaptation of immune responses of stickleback (*Gasterosteus aculeatus*) lake and river ecotypes. *Proceedings of the Royal Society* 274(1617): 1523-32

Scocchi, M., Pallavicini, A., Salagro, R., Bociek, K. and Gennaro, R. 2009. The salmonid cathelicidins: A gene family with highly varied C-terminal antimicrobial domains. *Comparative Biochemistry and Physiology, Part B* 152: 376-81

The Tree of Life Web Project. *The University of Arizona College of Agriculture and Life Science and The University of Arizona Library.* Website http://tolweb.org/tree/ 23.01.10

Tossi, A. 2005. Host defense peptides: roles and applications. *Current Protein and Peptide Science* 6(1): 1-3

Uzzell, T., Stolzenberg, E.D., Shinnar, A.E. and Zasloff, M. 2003. Hagfish intestinal antimicrobial peptides are ancient cathelicidins. *Peptides* 24(11): 1655-67

Zasloff, M. 2002. Antimicrobial peptides of multicellular organisms. *Nature* 415(6870): 389-95

Zhu, S. 2008. Did cathelicidins, a family of multifunctional host-defense peptides, arise from a cysteine protease inhibitor? *Trends in Microbiology* 16(8): 353-60

Zhu, S. and Gao, B. 2009. A fossil antibacterial peptide gives clues to structural diversity of cathelicidin-derived host defense peptides. *The Journal of the Federation of American Societies for Experimental Biology* 23(1): 13-20

Appendix A: Detailed methods

RNA extraction

- Sample is homogenized with 0.5 ml Trizol
- 0.5 ml Trizol added
- Incubated at RT for 5 min
- 0.1 ml BCP added and mixed for 15 sec
- Incubated at RT for 5 min
- Centrifuge at 12.000 x g for 15 min at 4°C
- Transfer aqueous upper phase to a new tube
- 0.5 ml isopropanol added and mixed for precipitation
- Incubate at RT for 10 min
- Centrifuge at 12.000 x g for 10 min at 4°C
- Pour supernatant off, careful not to loose the pellet
- Wash pellet with 1 ml 75% ethanol
- Centrifuge at 7.500 x g for 5 min at 4°C
- Pour ethanol off, careful not to loose the pellet
- Air dry the pellet until it is clear (10-15 min)
- Pellet resuspended in 40 µl RNase-free water

DNase I digestion

- Add 10 µg RNA
- Add 10 µl 10x DNase I buffer
- Add 2 µl DNase I
- Adjust volume to 100 µl with RNase-free water
- Incubate at 37°C for 10 min
- Incubate at 75°C for 10 min (DNase I inactivation)
- Cool on ice for 15 sec
- Add 250 µl cold 100% ethanol
- Centrifuge at 12.000 x g for 15 min at 4°C
- Discard supernatant and air-dry pellet
- Dissolve in 20 µl RNase-free water

DNA extraction

- To make 10 ml lyses buffer: 100 µl 1 M Tris at pH 8; 2 ml 0.5 M EDTA at pH 8; 8 ml H_2O and 250 µl 20% SDS. Add 5 µl RNase (20 mg/ml).
- Grind tissue with mortel in liquid nitrogen
- Add 1 ml of lyses buffer
- Incubate for 30-60 min at 37°C

- Add proteinase K (20 mg/ml) to final concentration of 100 µg/ml (5 µl to 1 ml)
- Mix gently
- Incubate at 57°C for 1-3 hours
- Cool to RT
- Add equal amount of sample and phenol/chloroform/isoamyl alcohol (25:24:1)
- Mix well
- Centrifuge at 12.000 x g for 15 min at 4°C
- Transfer upper phase to a new tube
- Add 80% of original amount of phenol/chloroform/isoamyl alcohol
- Mix well
- Centrifuge at 12.000 x g for 15 min at 4°C
- Transfer upper phase to a new tube
- Add 500 µl chloroform
- Mix well
- Centrifuge at 12.000 x g for 15 min at 4°C
- Transfer upper phase to a new tube
- Add 100 µl NaAce (3M) and mix
- Wash with 1 ml 100% cold ethanol
- Let stand on ice until the precipitate is visible
- Centrifuge at 5.000 x g for 10 min at 4°C
- Discard ethanol
- Wash pellet with 1 ml of 70% cold ethanol
- Centrifuge at 5.000 x g for 5 min at 4°C
- Discard ethanol and air-dry pellet
- Dissolve in 100 µl DNase free water

Nucleic acid quality control

- Measure nucleic acid concentration in sample, A260/A230, A260/A280 ratios and A340 using NanoDrop® Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA)
- A260/A280 indicates contamination with proteins and should be higher than 1.8
- A260/A230 indicates organic contamination (e.g. EDTA or phenol) and should be around 2.0
- A340 indicates undissolved particles and should be as low as possible

cDNA synthesis with First Strand cDNA synthesis kit #1612 from Fermentas

- Add the right volume of RNA for the final concentration of 1000 ng/reaction
- Add 1 µl of random hexamer primers
- Adjust volume to 12 µl with RNase-free water
- Mix gently and spin down in a micro centrifuge
- Incubate at 70°C for 5 min
- Place sample on ice

- Add 4 µl 5x reaction buffer, 1 µl RiboLock Ribonuclease Inhibitor and 2 µl 10 mM dNTP mix to each reaction
- Incubate at RT for 5 min
- Add 1 µl RevertAid M-MuLV Reverse Transcriptase

cDNA synthesis reaction in a 4-block PCR machine

1. 25°C for 10 min
2. 42°C for 1 hour
3. 70°C for 10 min
4. 12°C forever

PCR: volume for one sample

- Add 2 µl of 10x buffer
- Add 2 µl of 2.5 mM dNTP
- Add 0.4 µl Taq enzyme
- Add 0.8 µl of each primer
- Add 1 µl of template
- Adjust volume to 20 µl with RNase-free water

The general program for the PCR reactions:

1. 94°C for 5 min
2. X°C for 45 sec (annealing step)
3. 70°C for 1 min
4. 94°C for 45 sec
5. 72°C for 10 min
6. 12°C forever

Steps 2. -4. were repeated 33 times

Gel electrophoreses

- Mix 1.2 g of agarose with 100 ml 1x Tris-acetate-EDTA (TAE) buffer
- Boil to dissolve agarose
- Cool to 60°C
- Add 3 µl EtBr
- Pour the mixture into a sealed agarose apparatus and place a gel-comb to make wells.
- Let cool until gel is solid (approx. 30 min)
- Place in gel electrophoreses tank and add 1x TAE buffer
- Load Gene Ruler 100 bp ladder from Fermentas and samples into wells
- Run on 80 V for 45-60 min
- Analyze with UV light camera connected to a computer

Gel extraction

- Cut desired bands from gel
- Weigh the gel pieces with the bands
- Add 3 volumes of Gel solubilisation solution
- Incubate at 50-60°C for 10 min or until dissolved, vortex briefly 3-4x during incubation
- Add 500 µl column preparation solution to each binding column
- Centrifuge at 16.000 x g for 1 min and discard flow-through
- Add 1 volume of 100% isopropanol
- Mix well
- Add solution to a binding column
- Centrifuge at 16.000 x g for 1 min and discard flow-through
- Add 700 µl wash solution to column
- Centrifuge at 16.000 x g for 1 min and discard flow-through
- Centrifuge at 16.000 x g for 1 min again and discard flow-through
- Preheat elution solution to 65°C
- Add 50 µl of elution solution to the centre of the column
- Incubate for 1 min at RT
- Centrifuge at 16.000 x g for 1 min

Cloning

Preperation of agar plates and medium

Agar plates

- Mix 1 g yeast extract, 2 g Trypton, 2 g NaCl and 3 g Agar.
- Add to 200 ml distilled water
- Autoclave
- Let cool
- Add 200 µl spectinomycin
- Pour on plates and let cool until solid

Medium

- Mix 2.5 g yeast extract, 5 g Trypton and 5 g NaCl
- Add to 500 ml distilled water
- Autoclave

TopoTA cloning

- For 4 reactions add 1.32 µl salt; 1.68 µl H2O and 1 µl TOPO® vector in one tube
- Divide into four tubes
- Add 1 µl PCR reaction to each tube
- Incubate at RT for 5 min
- Put on ice

- Add 25 µl One Shot® chemically competent E.coli cells and mix together
- Incubated on ice for 10 min
- Heat-shock the cells at 42°C for 30 sec and put on ice
- Add 100 µl S.O.C. medium to each tube
- Incubate at 37°C for 1 hour with shaking
- Spread 10-50 µl of medium over LB agar plates with spectinomycin
- Keep at 37°C over night

Colony PCR

- Mix a basic PCR reaction master mix (Appendix A, section 4.1)
- Mix 3 ml LB medium with spectinomycin in sterile tubes
- Divide PCR reaction into PCR tubes
- Pick colonies with pipette tips
- Touch the PCR reactions with the tips
- Put the tips into corresponding tubes with the LB medium
- Incubate at 37°C with shaking, over night
- Use general program for PCR (Appendix A, section 4.2) where X is 60°C

Plasmid isolation

- Centrifuge 3 ml culture at 11.000 x g for 30 sec
- Discard the supernatant
- Add 250 µl Buffer A1 and vortex well
- Add 250 µl Buffer A2 and mix gently by inverting 6-8 times
- Incubate at RT for 5 min
- Add 300 µl Buffer A3 and mix by inverting 6-8 times
- Centrifuge at 11.000 x g at RT for 5 min
- Repeat last step if supernatant is not clear
- Place NucleoSpin Plasmid Column in a Collection Tube
- Pipette max 750 µl of supernatant onto column
- Centrifuge at 11.000 x g for 1 min
- Discard the flow-through
- Repeat from centrifuge until no supernatant is left
- Add 600 µl Buffer A4
- Centrifuge at 11.000 x g for 1 min
- Discard flow through
- Centrifuge at 11.000 x g for 2 min
- Discard the collection tube
- Elute DNA with 30 µl Buffer AE into micro centrifuge tube
- Incubate at RT for 1 min
- Centrifuge at 11.000 x g for 1 min

Cutting plasmids

- Add 12 µl H2O
- Add 2 µl Buffer 3 (New England Biolab, Ipswich, MA, USA)

- Add 5 µl DNA
- Add 0.3 µl EcoRI (New England Biolab, Ipswich, MA, USA)
- Mix carefully and centrifuge briefly
- Incubate at 37°C for 1 hour
- Run 10 µl on a agarose gel

Sequencing

- For one reaction:
- Add 1 µl DNA
- Add 1.5 µl 5x Buffer
- Add 1.6 µl primers (0.8 µl GW1 and 0.8 µl GW2 primers)
- Add 1 µl BigDye Terminator v3.1 (Applied Biosystems)
- Add water up to final volume of 10 µl

Program for sequencing

1. 96°C for 10 sec
2. 96°C for 10 sec
3. 50°C for 5 min
4. 60°C for 2 min
5. 12°C for 7 min
6. 12°C forever

Repeat steps 2. -4. 25 times

Precipitation of DNA for sequencing

- Glycogen mixture: 900 µl H2O, 100 µl 3M NaOAc and 5 µl frozen Glycogen (20 mg/µl)
- For each reaction: add 50 µl glycogen mixture and 125 µl ice cold 100% ethanol
- Centrifuge at 10.500 rpm for 15 min at 4°C
- Remove supernatant
- Wash with 250 µl 70% ethanol
- Centrifuge at 10.500 rpm for 5-7 min and remove ethanol
- Wash with 250 µl 70% ethanol
- Centrifuge at 10.500 rpm for 5-7 min
- Remove EtOH
- Air-dry in the dark for 10 min
- Add 10 µl HiDi
- Incubate at RT for a short time
- Vortex and spin down
- Moved to a sequencing plate and close with grey rubber lid
- Analyze in the ABI sequencer

Appendix B: Additional information on the degenerate primers

To construct the degenerate primers, two alignments were used:

a. The amino acid alignment that was published in 2009 (Scocchi et al. 2009).

b. An alignment was made between cathelicidin nucleotide sequences in rainbow trout (*Oncorhynchus mykiss)*, grayling (*Thymallus thymallus)*, arctic charr (*Salvelinus alpinus*) and cod (*Gadus morhua)* by using the Clustal W program, accessible on the European Bioinformatics Institute webpage. Shaded sequences equal the primer sequences, the red and grey shaded sequence is primer fDegF1.1 and the grey and yellow shaded sequence is primer fDegF1.2.

```
                                                                  fDegF1
rtCath2     AATCAGACATGAAGATGAAGGCTCTGGTGAGATCTCTCCTCTTGCTCGCTGTGGCTAACC 60
thyth       --------ATGAAGATGAAGGTCCAGGTGAGATCTCTCCTCCTGCTCGCTGTGGCTGGCC 52
rtCath1     ----AGACATGAAGATGAAGGTCCAGGTGAGATCTCTCATCCTGCTCGCTGTGGCTGTCC 56
charr       -------GTTCAAGATGAAGACTCAGATCCAATCCTACTTTTTGCTGTCGGTGGCCATCC 53
codCath1    ---------------GAATTGTTCAGCACAAAGCCAA------ACTGAAAATGACGACTC 39
                           *     * *     * *          **     ** *    *

rtCath2     TGCTGGTCA-GAGGTCAGACCCAGACTGAGACCAGGTATGAAGACATCATCACAGCTGCT 119
thyth       TGCTGGTCA-GAGGTCAGACCCAGACTGTGACCAGGTATGAAGACATCATCACAGCTGCT 111
rtCath1     TGCAGGTCA-GATCTCAGAACCAGACTGAGACCAGATATGAAGACATCATCTTAGTTGCT 115
charr       TGCTGGTGAAGGCCCCAGACACCATCTGAGCTCAGCTATGAGGACGTCTTATCAGGC-CT 112
codCath1    AGATGA----GATTGTTGTGCTTCGCTGCGGTTACCT-TACTGGCTGAGGCTCAGATGAT 94
                 *  *     *       *       *** *   *  * *   * *        **    *

                                        fDegF1.1-------fDegF1.2
rtCath2     TCAAATCAGCTGCTTCCTATGGAAGAGCAGGCTTTCCATCCGCTTCTGAACCAGCTGGAA 179
thyth       TCAACTCAGCTGCTTCCTGGGGAAGAGCAAGCTTTCCGTCCGCTTCTGAACCAGCTGGAA 171
rtCath1     TTGCCTCAGCTGCTTCCTGGGGAAGAGCAGGCTTTCCGTCCAATTCTGAACCAGCTCCAA 175
charr       GTTTGCCAGTGGAGA--TGCTGACCGACGCACTTTCCGTCCACTTCTAGAAGAACTCCAA 170
codCath1    CCCGGACCCCTTTATCTTTCCTCTGA--AAAACTTCCGCCCCCTGCTGGACCAGCTTCGG 152
                      *         *              ****  **  * **  *  * **

                                                        fDegF1.3
rtCath2     GTCGAGACTTTGAATACAGAGGATGTGGACCAGTCTGAGGTGTCTGT---AAGGTTGAGC 236
thyth       GTGGAGAATTTGAATACAGAGGATGTGGACCAGTCTGAGGTGTCTGT---AAGGTTGAGC 228
rtCath1     GTCAAGACTTTGAATACAGAGGATGTGGACCAGTCTGAGGTGTCTGT---AAGGCTGACC 232
charr       GAGGACACTTTGAACTCTGCAGA---GAA---------------------AAGGCTGAGT 206
codCath1    GTTGAAACTGTCTATCCAGAGGGTGTGGACCTGTCCACCATGTCAGTGAGAAAAATGACA 212
                *   * * * *  *  * *  *    * *                      **   ***

rtCath2     TTCCCCCTGCAGGAGACGCTCTGCAGTAAGGC--ACAGGGGCA-GCGAGGCCAACCATGC 293
thyth       TTCCCCCTGCAGGAGACTCTCTGCAGTAAGAC--ACAGGAGCA-GCAAGGCAAACCATGC 285
rtCath1     TTCCCCATACAGGAGACTTTCTGTAGTAAATC--ACAGGGGCA-GCCAGGCAAACCATGC 289
charr       TTCGTTGCAATGGAGACTGTGTGCGATGGGGC--AC------A-GCAAGGCATGCCATGC 257
codCath1    TTCCCTGCACAAGAGTTAGACTGCAGCCAAGTGAACACTTCCATGCCCGGACAGCAATGT 272
                ***         ***      **          **      * **  **    * ***

             fDegF2/fDegR0.5
rtCath2     CCTCTGAAGAAAAATGGGAAACGAATGATGTGCAGCATGGAGGTCAGACATCCGATTCTG 353
thyth       CCCCTGAAGAAAAATGGGAAACTAATTATGTGCAGCATGGAAGTCAGACATCCGATTCTG 345
rtCath1     CCTCTGAAGAAAAATGGGAAACTAATGATGTGCAGCATGAAGGTCAGACATCCGATTCTG 349
charr       CCTCTGAAGCAGAATGGGCAAGTATTGTCATGCAACCTAGCGGTTAGCCATGCAGTCCAG 317
codCath1    CCACTTAAGGAAAACGGGAAAATGATGAATTGCAATTTCACCCTGAGCTACATAAACCAA 332
                ** ** *** * ** *** **    *    ****   *     * **  *       *
```

```
rtCath2     GAGATTGGCAACACCCTCAACACTGACCGGTCTGACATTTCTTGTGAATACATGGAGGCA 413
thyth       GAGGCTAGCAACACCCTCAACACTGACCTGTCTGACATTTCTTGTGAATACATGGAGGCA 405
rtCath1     GAGGCAAGCAACAACCTGAACACTGACCTGTCTACATTTGTTTGTGAATACATGGACGCA 409
charr       GATGCCAGCAACA---------TTGATACCTCCAAAATGTCTTGTCAGCCCATGGAGAAA 368
codCath1    GATGCTGACA------------TTCAAGGCTTTCAGTTCAACTGTGATGCTGCAATCAAG 380
                ** *      **            * *    *      *    *** *

                                fDegF3/fDegR1
rtCath2     GAAGTT---------CAGAAGATTCGGACAAGAAGAGGCAAGGACAGCGGAGGCCCTAAG 464
thyth       GAAGACGCTATGCAGCAGAAGTTTCGGACAAGAAGAAGCAAGTCCAGC------TCTAAT 459
rtCath1     GAAGATGCTTTGCAGCAGAAGATTCGGACAAGAAGAAGCAAAGTCAGAATATGCTCCAGA 469
charr       GAGGAGGCTTTAAAGCAGAAGATTAGGATGAGGAGGAGTAAGTCTGGAAAGGGTTCCAGC 428
codCath1    GAGGCAACCCT----------GACCAGAGTGAGAAGGAGCAGGTCTGGCCGTGGATCAGGA 431
                ** *                *    *    ** **  * *      *        *

rtCath2     ATGGGTAGAAAAG------ATTCCAAGGGGGGTTGGAGA---GGTCGTCCTGGGAGTGGC 515
thyth       GGGGGTAGAAAAG------GTTCCAAGGGCGGTAGCAAA---GGTCGTCCTGGGAGTGGC 510
rtCath1     GGCAAAAATTGTGTCTCTCGTCCTGGGGTTGGCTCCATAATTGGTCGTCCTGGGGGTGGC 529
charr       TCCAAGGGATCCAAG---GGATCCAAGGGGGGTTCC---------CGTCCTGGGGGTGGA 476
codCath1    AAAGGGGGCCGTGGC---GGATCTAGGGGGTCCAGTGGGTCCAGAGGCTCCAAG---GGA 485
```

When constructing degenerate primers from amino acid sequences, codon bias for the species of interest must be accounted for. In stickleback only two codons had a frequency under 5%, TTA and CTA, which both code for the amino acid leucine. Therefore it was possible to exclude these codons from the primer construct.

The next step in degenerate primer construction is to assign symbols that correctly represent the option between multiple bases. The wobble IUPAC-IUB symbols were used for this. They are **R** (for A or G bases), **Y** (C or T), **M** (A or C), **K** (G or T), **S** (G or C), **W** (A or T), **H** (A or C or T), **B** (G or T or C), **V** (G or C or A), **D** (G or T or A), **N** (G or A or T or C).

The final step of construction is to order the primers, which was done from Tag Copenhagen A/S.

Table B-1 ***Overview of ordered degenerate primers.***

Name	Degenerate sequence	Tm (°C)		Result			Species	
		Tm1	*Tm2*	*only F*	*Both*	*only R*	*Tried*	*Positive*
fDegF1	5'- CTS KCK GTG GCY RDC CTG CWN GT –3'	67,5	54	+	+	-	SB,AS	SB,AS
fDegF1.1	5'- CRV RCT TTC CRY CCV MTK CT -3'	60,0	54	-	+	-	SB,AS	AS
fDegF1.2	5'- TTC CRY CCV MTK CTG RAM SA -3'	58,7	54	-	+	-	SB,AS	AS
fDegF1.2b	5'-TTC CRY CCV MTK CTG RAC CA-3'	59,7	54	-	+	-	SB,AS,Cod	AS,Cod
fDegF1.3	5'-CWN TCY RMS RTG TCW GTR AG-3'	57,3	54	-	+	-	SB,AS,Cod	AS
fDegF2	5'- TGC CCY CTG AAR MAR AAY GG -3'	58,3	54	-	+	-	SB,AS	SB,AS
fDegF2b	5'-TGY CCH CTK AAG RAA AAY GG-3'	55,9	54	-	-	-	SB,AS,Cod	none
fDegF3	5'- CAR AAR WTT MGG AYR AGR AG -3'	53,2	NA	-	NA	-	NA	NA
fDegF3b	5' - CAG AAG WYY MGR RGR AGA AG-3'	57,3	NA	-	NA	-	NA	NA
fDegR0.5	5'-CCC RTT YTK YTT CAG RGG YC-3'	59,4	54	-	+	-	SB,AS,Cod	AS
fDegR0.5b	5'-CCC RTT TTY CTT MAG DGG RC-3'	58,0	54	-	+	-	SB,AS,Cod	AS,Cod
fDegR1	5'- CTY CTY RTC CKA AWY TTY TG -3'	53,2	54	-	+	-	SB,AS	SB,AS
fDegR1b	5'-YCT TCT YRY YCK RRW CTT CT-3'	55,3	54	-	-	-	SB,AS,Cod	none

Tm1 is the melting temperature given by Tag Copenhagen, Tm2 is the best observed temperature when gradient PCR was performed. The results column indicates if the primer gives bands with an appropriate forward and reverse (Both) or on its own (only F, only R). Only the forward (F) primer fDegF1 and no reverse (R) primer were found to amplify on its own. The species column indicates which species the primer anneals to of the once that were tried. Primers fDegF3 and fDegF3b were designed for 3'RACE procedure and therefore these primers could not be used for normal PCR since there was no designed reverse primer downstream of the sequence (see Clustal W alignment).