

Master of Science Thesis

Cloning and Sequencing of G Protein encoding genes of *Balanus improvisus*



Supervisor Ulrika Lind

Examiner Anders Blomberg

Department of Cell and Molecular Biology
University Of Gothenburg

Author SABIHA AMIR

Department of Chemical and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2010.

The Author grants to Chalmers University of Technology the non-exclusive right to publish the Work electronically and in a non-commercial purpose make it accessible on the Internet. The Author warrants that he/she is the author to the Work, and warrants that the Work does not contain text, pictures or other material that violates copyright law.

The Author shall, when transferring the rights of the Work to a third party (for example a publisher or a company), acknowledge the third party about this agreement. If the Author has signed a copyright agreement with a third party regarding the Work, the Author warrants hereby that he/she has obtained any necessary permission from this third party to let Chalmers University of Technology store the Work electronically and make it accessible on the Internet.

Cloning and Sequencing of G Protein alpha encoding genes of *Balanus improvisus*
SABIHA AMIR

© SABIHA AMIR, 2010
Examinar: Prof. Anders blomberg

Technical report no 2010:xx
Department of Chemical and Biological Engineering
Chalmers University of Technology
SE-41296 Göteborg
Sweden
Telephone +46 (0) 31-772 1000

Cover: Picture shows the adult form of the *Balanus improvisus* (Photo by Sergej Olenin)

I dedicate this little effort of mine to my parents for their never ending moral support, encouragement and prayers. Specially dedicated to my husband for his cooperation, financial support and love that enabled me to achieve this goal.

Contents

ABSTRACT.....	1
LIST OF ABBREVIATIONS.....	2
AIM.....	3
INTRODUCTION	3
Barnacles and biofouling	3
Adhesion of barnacles.....	4
Medetomidine and barnacles	4
G protein coupled receptors (GPCRs)	4
Heterotrimeric G protein.....	4
G protein and barnacles.....	6
MATERIAL AND METHODS	7
cDNA Synthesis.....	7
RLM- Race.....	7
PCR amplification.....	7
Gel Electrophoresis of DNA	8
Purification and digestion of PCR product	8
TOPO Cloning of $G\alpha_q$ and $G\alpha_s$ PCR products.....	8
G-protein expressed sequence tags (ESTs).....	8
Designing Gene Specific Primers (GSPs).....	8
RESULTS	10
DISCUSSION	21
ACKNOWLEDGEMENTS.....	23
REFERENCES	24

ABSTRACT

The barnacle *Balanus improvisus* is a major fouling organism in Swedish waters. Settling of barnacles onto ship hulls leads to biofouling. Antifouling paints are used to prevent the settlement of the barnacles on ship hulls, but some of these have severe impact on marine ecosystems. G-protein-coupled octopamine receptors of *B. improvisus* have already been cloned and sequenced. It was discovered that these octopamine receptors are activated by the novel antifouling agent medetomidine, leading to inhibition of the settling process of *B. improvisus*. Research on *B. improvisus* has been conducted to provide alternatives to toxic substances such as tributyl tin (TBT), copper metal oxides or organic biocides in marine paints.

The knowledge about G proteins of barnacles is very limited. We have cloned and sequenced two G protein subunits; $G\alpha_q$ and $G\alpha_s$. Cloning of $G\alpha_q$ and $G\alpha_s$ was performed by PCR amplification using sequence information based on ESTs from a *B. improvisus* cDNA library. cDNA was used as template. The obtained sequences were aligned with lobster *Homarus americanus* sequence. In $G\alpha_q$ different variants of 5' ends were found whereas in $G\alpha_s$ different variants at the 3' end are present. Interestingly $G\alpha_s$ in *B. improvisus* has 165 nucleotides longer open reading frame as compared to the one in *H. americanus*.

This is a first step in the cloning of G-protein subunits of *B. improvisus*. Functional studies of G-protein with GPCRs will be performed in the future. The enhanced knowledge of G-protein and GPCRs in barnacles will contribute to the development of more specific antifouling paints.

LIST OF ABBREVIATIONS

GPCRs	G -protein Coupled R eceptors.
GDP	G uanosine D iphosphate
GTP	G uanosine T riphosphate
RT-PCR	R everse T ranscription P olymerase C hain R eaction
DMSO	D imethyl S ulfoxide
cPCR	C olony- P olymerase C hain R eaction
RLM-RACE	R NA L igase M ediated R apid A mplification of c DNA E nds
EST	E xpressed S equence T ags
GSP	G ene S pecific P rimer

AIM

The aim of this project is to clone and sequence different subunits of $G\alpha$ proteins from the barnacle, *B. improvisus*.

This report focuses on:

Cloning and sequencing of the two G protein subunits $G\alpha_s$ and $G\alpha_q$.

INTRODUCTION

Barnacles and biofouling

Biofouling is a worldwide problem in marine systems and barnacles are the most problematic biofouling marine organisms. Biofouling results in an increase in drag, which in turn leads to increased fuel consumption, hull cleaning and removal of paint (Callow and Callow, 2002). *B. improvisus* is the barnacle species causing the most severe fouling problems in the Baltic sea (Lind et al., 2010). The life cycle of *B. improvisus* has seven larval stages, and between the sixth and seventh transitions, nauplii larva transforms into cyprid larva. In order to complete the transition to adult life, cyprid larvae must attach to a hard substrate (Fig.1).

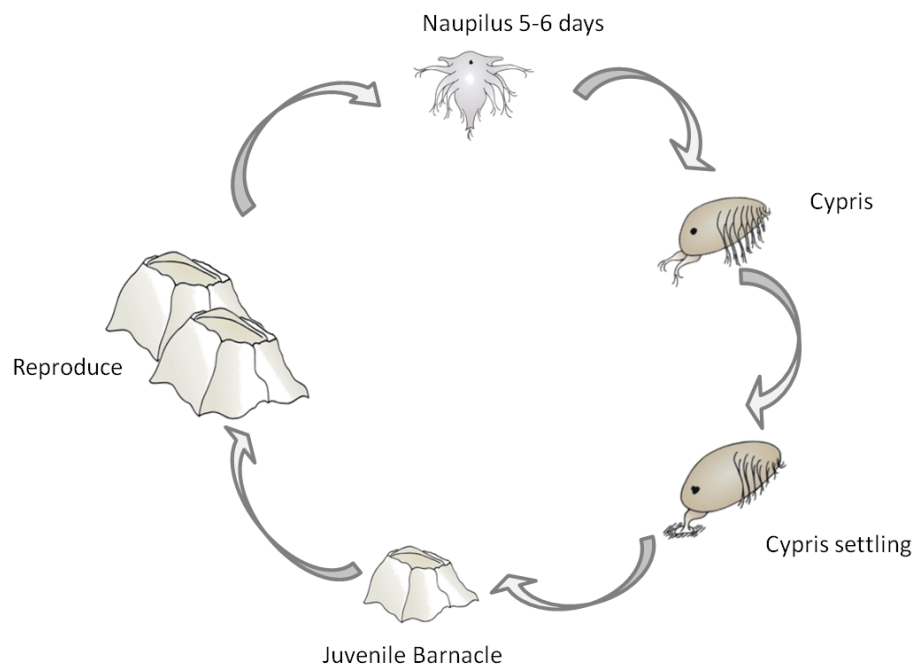


Figure 1: The Life Cycle of *B. improvisus*
drawing by Susanna Falkbring

Adhesion of barnacles

After identifying a suitable surface to adhere cyprid larva secretes adhesive proteins by a pair of cement glands. The cement glands are connected by cement ducts, which widens into a muscular sac. The muscular sac is connected to an antenna by another cement duct and pours proteinaceous cement into antenna. Cyprid cement flows around and embeds the attachment organs and the cyprid larva is able to attach itself to the surface and metamorphoses into a calcified adult barnacle (Odling et al., 2006).

Medetomidine and barnacles

Traditional antifouling substances like tin and copper are highly toxic to non-target organisms in the marine environment. Organotin compounds such as TBT are now banned worldwide. Medetomidine is now being used as effective alternative in marine paints (Krang and Dahlstrom, 2006). Since the 1980's, medetomidine has been used as a sedative agent in veterinary medicine. Medetomidine functions as α_2 -adrenoceptor agonist and binds to G protein coupled octopamine receptors in invertebrates and α_2 -adrenoceptor in vertebrates. Medetomidine strongly inhibits the settling process of the barnacles *B. improvisus* at very low (nanomolar) concentrations (Lennquist et al., 2010).

G protein coupled receptors (GPCRs)

G protein coupled receptors represent one of the largest and most diverse groups of proteins present in most eukaryotes. All GPCRs have seven transmembrane α helical regions, an extracellular N-terminus and an intracellular C terminus. GPCRs bind with a wide range of ligands that modulate vital physiological events like neurotransmission, sensory perception, chemotaxis and cell communication (Tuteja, 2009). The ligands of GPCRs can be various types of substrates like amino acids, proteins, biogenic amines, nucleotides, lipids and Ca^{+2} ions. When activated by ligands, G protein coupled receptors undergo a conformational change that leads to G protein activation by promoting the exchange of GDP/GTP associated with the $\text{G}\alpha$ subunit (Oldham and Hamm, 2008).

Heterotrimeric G protein

Heterotrimeric G proteins are molecular switches that turn on signaling processes for hormones, neurotransmitters, and photons. The G proteins are composed of three subunits, α , β and γ . G proteins exist as trimers with GDP bound to the α subunit in its inactive form. Receptor activation causes G proteins to dissociate into $\text{G}\alpha$ and $\text{G}\beta\gamma$ because GDP is replaced by GTP (Xu et al., 1998). The switching function of the G protein depends upon the identity of the α subunit; four $\text{G}\alpha$ subunit classes ($\text{G}\alpha_s$, $\text{G}\alpha_i$, $\text{G}\alpha_q$, and $\text{G}\alpha_{12}$) have been proposed based on amino acid sequence analysis, each performing specific functions (Han and Watson, 2005). The $\text{G}\alpha_s$ activates an enzyme called adenylyl cyclase which in turn catalyses the synthesis of the second messenger cAMP. In contrast, $\text{G}\alpha_i$ inhibits the activity of adenylyl cyclase. The $\text{G}\alpha_q$ activates phospholipase C that catalyses the synthesis of two second messengers (DAG and IP_3). The $\text{G}\alpha_{12}$ proteins regulate cytoskeletal assembly (Lodish, 2008). Within the adrenergic receptor family for

example, the receptors couple to different G-proteins. β -adrenergic receptors couples to $G\alpha_s$, the α_1 -adrenergic receptor couples to $G\alpha_q$ and the α_2 -adrenergic receptor is coupled to $G\alpha_i$.

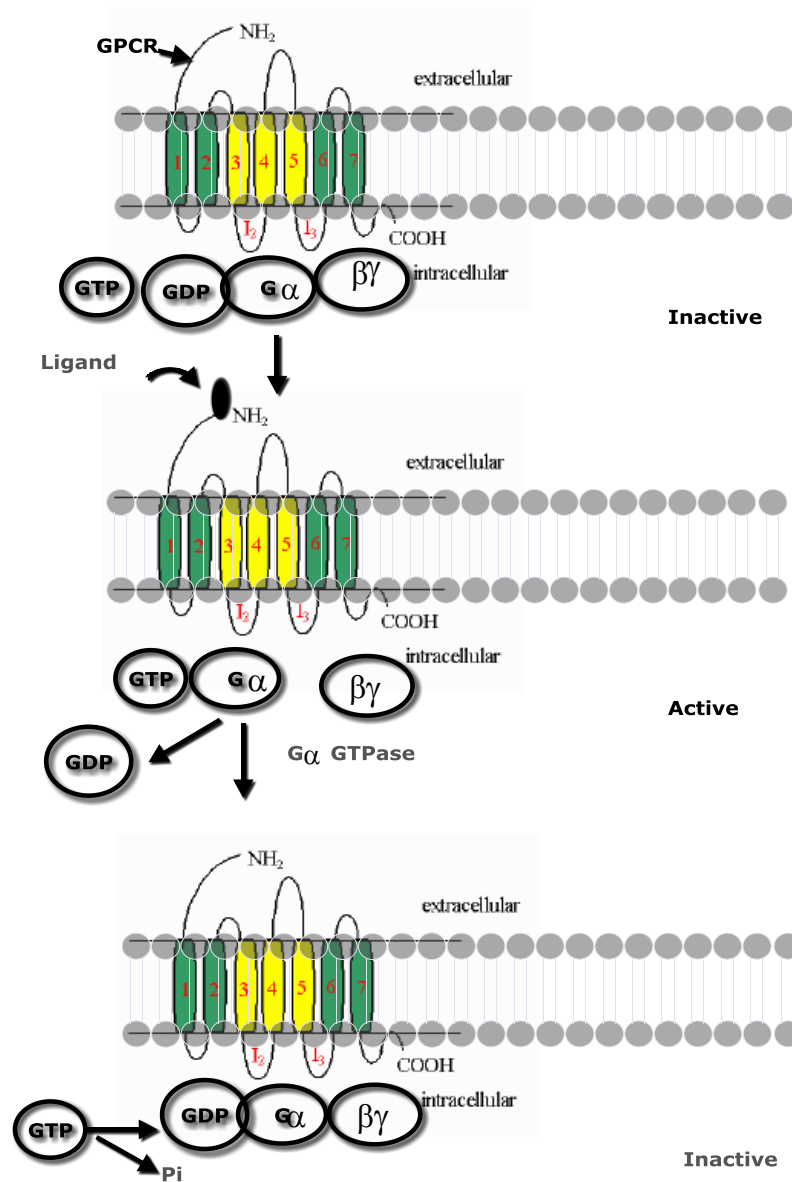


Figure 2: Interaction of GPCR and G protein

G protein and barnacles

Five different G protein coupled octopamine receptors, one α and four β -like receptors, have been recently cloned and characterized by Lind and Blomberg at the University of Gothenburg (Lind et al., 2010). Before this only two GPCRs of the barnacle had been cloned and sequenced (Isoai et al., 1996) (Kawahara et al., 1997). The interaction of the novel antifouling agent medetomidine with the cloned octopamine receptors was studied. All five receptors increase cAMP signaling in the cell, however the main signaling pathway for the α like receptor was via increasing Ca^{+2} levels in the cell. In order to produce more specific antifouling paints it is important to know more about the GPCRs and G-proteins of the barnacles. More knowledge about G-proteins of barnacles and their interactions with GPCRs will help to design non toxic substances. This work focuses on the cloning and sequencing of G-protein encoding genes of *B. improvisus*. Interaction between G-proteins and GPCRs of the barnacles will be studied in the future.

MATERIAL AND METHODS

cDNA Synthesis

cDNA was synthesized from 1 µg of total RNA from cyprid by following the protocol of Superscripts™ III First-strand synthesis system for RT-PCR (Invitrogen life technologies). RNA together with 1 µl of oligo dT and 1 µl of 10 mM dNTP was heated to 65 °C and then incubated on ice for 1 min. Then cDNA mix was added and incubated in 50 °C for 50 min.

cDNA synthesis mix

10XRT Buffer	2 µl
25 mM MgCl ₂	4 µl
0.1 M DDT	2 µl
RNase OUT™	1 µl
Super Script™ III RT	1 µl

The cDNA synthesis was terminated by incubating at 85°C for 5 min. The cDNA was then treated with 1 µl RNase H and incubated at 37°C for the removal of RNA.

RLM- Race

RNA ligase mediated rapid amplification of 5' and 3' cDNA ends was done by following Gene Racer™ kit (Invitrogen life technologies). In the first step total RNA was treated with calf intestinal phosphatase (CIP) to dephosphorylate non-mRNA or truncated mRNA. Dephosphorylated mRNA was then precipitated and the mRNA cap structure was removed. An RNA Oligo was ligated to decapped mRNA following the kit protocol. mRNA was reverse transcribed into cDNA using Gene Racer Oligo dT primers. For the amplification of the 5' end of cDNA Gene Racer 5' primer and Nested 5' primer together with gene specific primers were used. The amplification of 3' end was done by using Gene Racer 3' primer and Nested 3' primer with gene specific primers.

PCR amplification

For the cloning of $G\alpha_s$ and $G\alpha_q$ genes in *B. improvisus*, cDNA was used as template in PCR amplification by using different primer combinations based on cDNA library of *B. improvisus* made by M. Alm Rosenblad and A. Blomberg.

The PCR reactions were performed by preparing 50 µl of Pfu Ultra enzyme reaction mix in the following concentrations; 1 µl of cDNA was added to 49 µl of reaction mix containing 1x Buffer, dNTP (0.2 mM), DMSO (5%), primers (fw and rev) 0.5 µM, and Pfu Ultra enzyme 1 µl.

Master mix for cPCR

For colony PCR 19 μ l of master mix was prepared for each reaction containing the following. M_{13} Fw, and M_{13} Rev (0.5 μ M each), 1x Buffer true start, dNTP (0.2 mM), DMSO 5%, $MgCl_2$ (1.5 mM), and 0.20 μ l of True start Taq Polymerase. 1 μ l cell suspension was then added to obtain final volume of 20 μ l for each reaction.

Gel Electrophoresis of DNA

Agarose gel electrophoresis was used to analyze the sizes of DNA fragments. In order to prepare 1% agarose gel to separate bigger fragments of more than 1000 bp, 1g of gel was weighed and melted in 100 ml 1X TAE buffer. For smaller fragments 1.5% of gel was formed by adding 1.5g of gel into 100 ml of buffer. Ethidium bromide was added to the gel to enable the visualization of DNA using UV light. DNA samples were prepared by mixing DNA solution with gel loading dye (Fermentas). Gel electrophoresis of DNA was performed with the constant voltage of 100 volts.

Purification and digestion of PCR product

After separating PCR products on gel, they were purified with a gel extraction kit (Qiagen) following the manufacturer's instructions. Shortly, the gel was solubilized by adding 3 volumes of buffer QG and incubated at 50°C for 50 min. The sample was added to the QIA quick column and the DNA was washed by buffer PE and finally eluted in 50 μ l elution buffer. In some cases purified PCR product was cleaved with the restriction enzyme Eco RI. 2 μ l of 10x buffer was mixed with 1 μ l of DNA then 16 μ l of H_2O and 1 μ l of Eco RI was added and incubated at 37°C for 1 hour.

TOPO Cloning of $G\alpha_q$ and $G\alpha_s$ PCR products

The PCR products were cloned into plasmids using the Zero Blunt® TOPO® PCR Cloning Kit for Sequencing (Invitrogen). 4 μ l of PCR product and 1 μ l of salt solution were added to 1 μ l TOPO vector, mixed and incubated for 30 minutes. Plasmids containing the ligated PCR product were transformed to *E. coli* by following the One Shot Chemical Transformation Protocol (Invitrogen life technologies). Shortly, cells were heat shocked for 30 s at 42°C, incubated at 37°C in SOC medium and then spread on ampicillin containing plates and incubated at 37°C overnight. Colonies were picked to perform colony PCR (cPCR). Positives colonies from the cPCR were cultured in 3 ml LB overnight and the plasmid was purified by following the QIA miniprep Kit Protocol (Qiagen). Plasmid concentration was determined using a Nano drop method, and purified DNA was sent to MWG biotech Germany for Sanger sequencing.

Expressed sequence tags (ESTs) Library

RNA was prepared from approximately 1000 cyprids. cDNA was prepared from RNA and then sequenced.

Designing Gene Specific Primers (GSPs)

Several GSPs were designed based on the *B. improvisus* ESTs library. Primer optimization was done by using different parameters of primer design like length, T_m and GC content (%).

Table 1. Primers sequence used for cloning

Primer Name	Sequence	Number of Base Pairs	Tm (°C)
Gaq_fl_fw	AGCCTCACCAGAAACCTTCA	20	60
Gaq_fl_rev	CAACGGAATGGTGTGTGAAA	20	56
Gaq_gap_fw	CTATGAACGGCGAAACGAGT	20	60
Gaq_gap_rev	ACGTTCTCGAAGCAGTGGAT	20	60
Gaq1fw_fl	TGGTGGAAATGCCGTGGAACACTAT	25	74
Gaq1_rev_fl	CGCCGTCACCGTCACCTCACTC	22	74
Gaq_fl_2fw	AAGTTTTGGCGAGTAGCAGAAAG	24	64
Gaq_fl_2rev	GTCACCGTCACCGTCTCACT	20	64
Gaq_fl_3fw	TTTAGAAAGAAACGGAAGTGCTG	23	64
Gaq_fl_3rev	GACTCGGCAATGACTAGCAG	20	62
Gaq_fl_4fw	CACAATCGGCAACAATACCA	20	58
Gaq_fl_4rev	GTCACCGTCACCGTCTCA	28	58
Gaq_fw5	TGGCCATGCAGAGTATGGTCAATGC	25	76
Gaq_rew5	GATGGGACGTCATGATCTTCTCCTC	25	76
Gaq_fw6	GCACCGCATCAATCGTGAAATTGAGA	26	76
Gaq_rew6	GCGGCGAACACGAAGCGGATGTT	23	74
Gaq_gapfwb	CGTGACGTCCATCATCTTCCTGGT	24	74
Gaq_gap_rewb	CTGGTACTCGTTTCGCCGTTTCATAG	25	76
Gaq_rew1b	CCGCCAACGGAATGGTGTGTGAAA	24	74
Gaq_fw3b	TAGAAAGAAACGGAAGTGCTGCTCGT	26	76
Gaq_gap_rev3b	CCGCGACTCGGCAATGACTAGCA	18	74
Gaq_fw4b	GATCCACAATCGGCAACAATACCAC	25	74
race_5'_Gaq_1	AGTTCCCTCCTAGCATCTCTCTTGT	25	74
race_5'_Gaq_2	GATGAATGTTGACTTCCCAGATTAC	26	74
race_5'_Gaq_3	AAGCCACGGCGATCCTCATCTGTGT	25	78
Gas_fw2	GGACTGCAGGCCCTATGACATGG	23	74
Gas_fw1	GGCAGTGCTTGCGATTTGTATGGTG	25	76
Gas_rew1	CCCACGTCGAACATGTGAAAGTTGA	25	74
Gas_2rev_fl	CGCTCGCTCGCTCTCTCACGAC	22	74
Gas_3rev_fl	GTGAAGTATTTGACCCTTTATCTCCC	26	74
race_3`_Gas_1	ACAAAACGAGCACCATTCGACAGGA	25	74
race_3`_Gas_2	CTACTCACCCACAGAACAGGATATTC	26	76
race_3_Gas_3	AGTCCTCACGTCCGGCATATTCTGA	24	74

RESULTS

To clone and sequence G protein encoding genes of the *B. improvisus* cDNA was used as template. PCR amplification was done by using different primer combinations based on ESTs (expressed sequence tags) from a *B. improvisus* cDNA library constructed by M. Alm Rosenblad and A. Blomberg (unpublished data).

There were two ESTs for $G\alpha_q$ in the EST cDNA library with a gap in between. To cover this gap, two primers were designed on either side of the gap, called gap fw and gap rev. PCR was performed and a band of expected size, about 300 bp was obtained (Fig. 3).

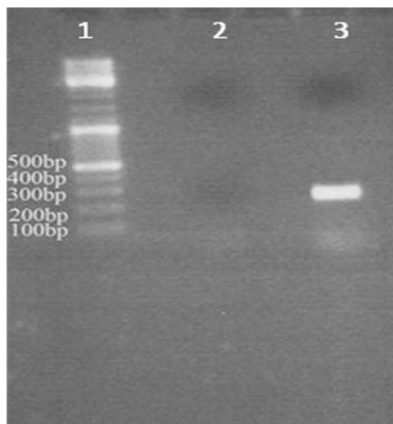


Fig. 3. $G\alpha_q$ fragment obtained from PCR. Sample (1) 100bp DNA ladder. An approximately 300bp band was obtained (well 3) covering a sequence gap of aligned ESTs. The following PCR program was used: 3 min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 52° C annealing for 30 s, 72° C extension for 2 min and final extension at 72° C for 10 min.

After performing gel extraction of the obtained DNA band it was ligated into the plasmid pCR[®] 4Blunt TOPO. The ligated fragment was transformed into *E. coli* and DNA was purified and sent for sequencing. In an attempt to amplify the $G\alpha_q$ full length sequence, different primers were designed located before the start ATG and after the stop codon (Fig. 4A).

A



B

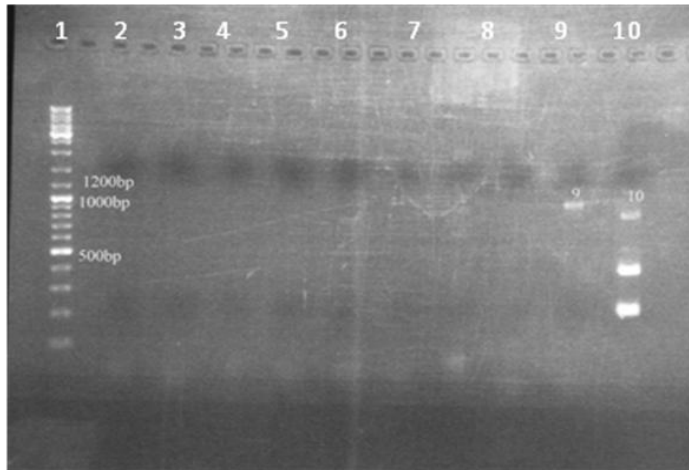
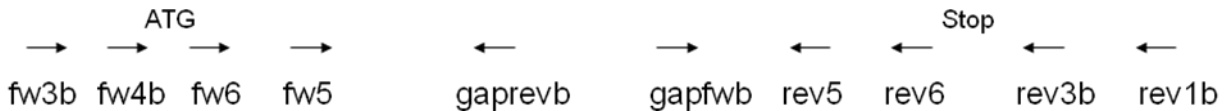


Fig. 4. PCR was performed by using different full length primers for $G\alpha_q$. A. Location of the primers. B. Bands were obtained only with gap fw, rev 1 (well 9) and with gap fw, gap rev (well 10). The following PCR program was used: 2 min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 49° C annealing for 30 s, 72° C extension for 2 min and final extension at 72 °C for 10 min.

The full length sequence of $G\alpha_q$ was not amplified. Only a smaller fragment of $G\alpha_q$ from the gap to the 3' end was amplified which was approximately 1000 bp in length (Fig. 4B).

A



In a second attempt to amplify the full length sequence of $G\alpha_q$ new primers were designed with a higher annealing temperature (Fig. 5A). Eight different primer combinations covering different parts of the coding region were used, but only four bands were obtained. The full length sequence of $G\alpha_q$ was not amplified, but all the $G\alpha_q$ coding region from gap to the 3' end was amplified (Fig. 5B).

B

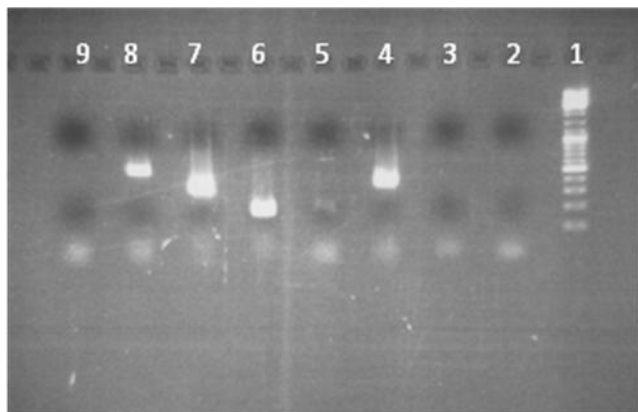


Fig. 5. Bands from fw6 and gap rev b, gap fw b and rev 5, gap forward b and rev 6, gap fw b and rev 3b were obtained. The following PCR program was used: 2min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 60° C annealing for 30 s, 72° C extension for 2 min and final extension at 72° C for 10 min).

5' Race for $G\alpha_q$

Our primers are based on ESTs from a cDNA library. In order to confirm that these EST sequences are correct, Race 5' $G\alpha_q$ primers were designed to amplify the complete 5' end of $G\alpha_q$ down to just below the start ATG. Touchdown and nested PCR were performed by using these primers together with Gene race primer (RLM-RACE kit, Invitrogen). RLM-RACE cDNA was used as template instead of ordinary cDNA. Touchdown PCR is a variant of PCR that aims to reduce non specific primer annealing by gradually lowering the annealing temperature. Nested PCR increases the specificity of PCR reaction. Two sets of primers are used in two successive PCR reactions.

The results of the touchdown PCR showed several faint bands, but after nested PCR a major band product of about 500 bp was obtained (Fig. 6). This PCR product was purified, cloned and sent for sequencing.

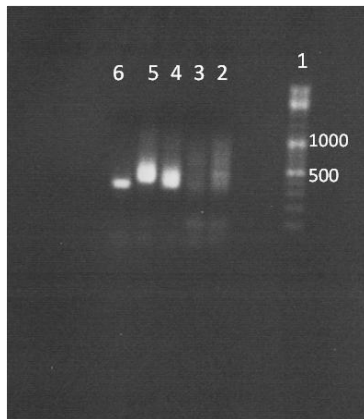


Fig. 6. 5' RACE for $G\alpha_q$. A distinct band was obtained in well no.5 (band size between 400 and 500 bp). The PCR program used was: 2 min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 65° C annealing for 30 s, 72° C extension for 1 min and final extension at 72° C for 10 min).

An EST from the cDNA library was aligned with the 5' Race sequences, both were similar. Thus sequence errors in ESTs could not explain the failure to obtain the full length $G\alpha_q$, when using 5' primers based on the ESTs.

3' Race for $G\alpha_q$

In order to investigate if it was possible to amplify the sequence from just below the ATG to 3' end of $G\alpha_q$, 3' Race was performed. Fw4b and the 3' Gene racer primer was used. Touchdown PCR and a subsequent nested PCR was performed with fw 6 or fw 5 combined with the nested primer from the RLM-RACE kit. Two fragments were obtained from 3' race which covers almost all the coding region and un-translated 3' end of the $G\alpha_q$ (Fig. 7).

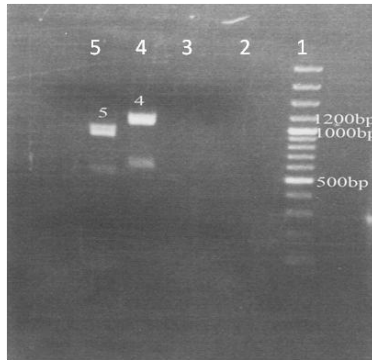


Fig. 7. Touchdown and nested PCR to amplify the 3' end of $G\alpha_q$. Two bands were obtained from nested PCR by primer combination fw 6 with 3' nested Gene racer primer (sample no.4) and fw 5 with 3' Gene racer primer from kit (sample no.5). PCR was run according to the following: 2 min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 58° C annealing for 30 s, 72° C extension for 2 min and final extension at 72° C for 10 min).

Amplification of $G\alpha_q$ full length sequence

To amplify the complete sequence of $G\alpha_q$ 5', Gene racer primer and nested 5' primer were used with gene specific primers (GSPs) rev 1b and rev 3b, respectively. A smear of unspecific bands was obtained, but a band with expected size of about 1400 to 1500 bp in length was also obtained (Fig. 8). This band covers the complete 5' end sequence and down to below the stop codon. The full length sequence of $G\alpha_q$ was thus amplified. By comparing the extreme 5' end with the previous 5' race short fragment and the ESTs sequences, it was found that they were not similar (Fig. 9).

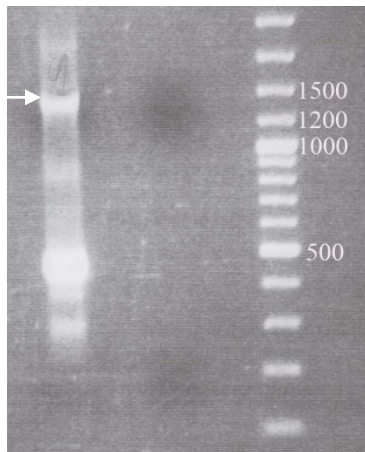


Fig. 8. $G\alpha_q$ 5' RACE PCR product. Running a 5' race with primers at the extreme 5' end and just below the stop codon, a band obtained which is about 1400 to 1500 bp (arrow indicates right band). The following PCR program was used: 2 min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 58° C annealing for 30 s, 72° C extension for 2 min and final extension at 72° C for 10 min.

Sequence alignment of *B. improvisus* $G\alpha_q$ 5' race short fragment with $G\alpha_q$ full length Sequence

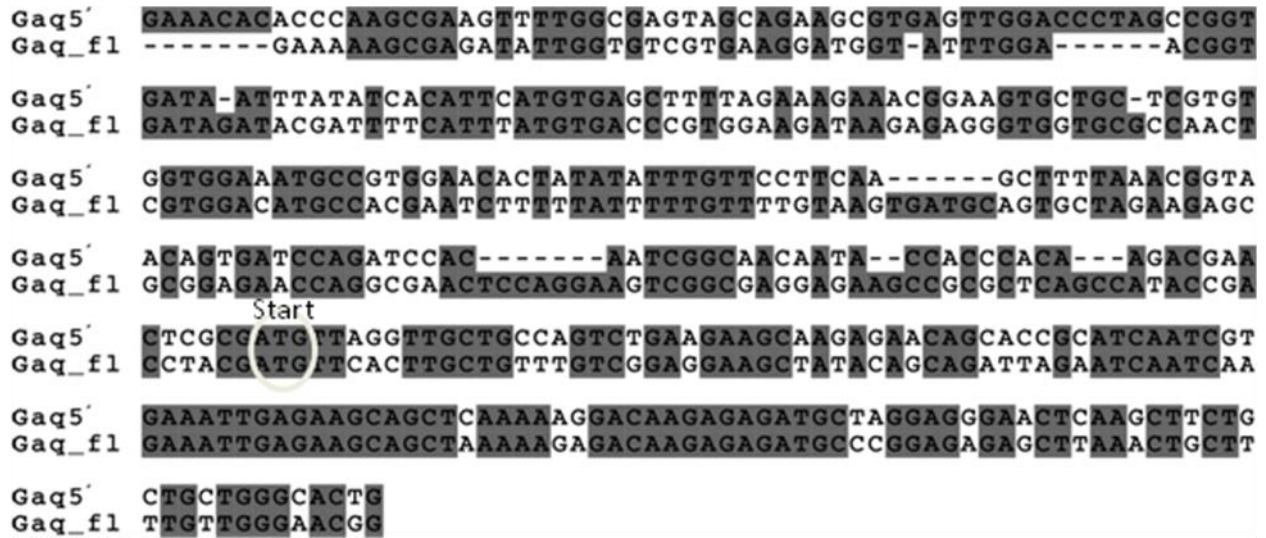


Fig. 9. Comparison of the 5' ends of the $G\alpha_q$ 5' RACE short fragment with the $G\alpha_q$ full length sequence obtained from 5' RACE. Nucleotide sequence alignment of *B. improvisus* $G\alpha_q$ was performed by using ClustalW program. It was found that the 5' end of the $G\alpha_q$ short fragment is different from the $G\alpha_q$ full length 5' end. Grey color indicates similar sequences while white indicates sequences are not similar.

Sequence alignment of *B. improvisus* $G\alpha_q$ and *Homarus* $G\alpha_q$

To compare our $G\alpha_q$ obtained sequences with other crustaceans, the sequence alignment of $G\alpha_q$ full length with *H. americanus* was done. The nucleotide sequence of the *B. improvisus* $G\alpha_q$ 5' end was aligned with *H. americanus*, the sequences were relatively similar (Fig.10 a). The nucleotide sequences of $G\alpha_q$ full length were translated and were aligned with amino acids sequences of *H. americanus*. Sequence alignment of $G\alpha_q$ full length and *H. americanus* shows very high sequence similarity (Fig. 10 b).

A

```

Homom_gaq      -----CCGTTGCTGTGCGGCG
Bam_5?gaq_fl   GAAAAAGCGAGATATTGGTGTCTGTAAGGATGGTATTTGGAACGGTGATAGATACGATTT

Homom_gaq      GCACATATATGACCTGTGTGA-----AGGAAATAACCCAAAAATT-GAGAAAAATTACAC
Bam_5?gaq_fl   TCATTTATGTGACCCGTGGAAAGATAAGAGAGGGTGGTGCGCCAACTCGTGGACATGCCAC

Homom_gaq      GATTACAGCCACTTTTCCCTTGGTGGAGAAAGCGGGGAAAGAGAAGAGGAAGAAGGAT
Bam_5?gaq_fl   GAAT-CTTTTTATTTTTGTTTG--TAAGTGATGCAGTGCTAGAAGAGCGCG-GAGAACC

Homom_gaq      AAGTTTATGTAAGGAGATTGTGGTGAAAGTCGTCAAAATTAACGTTATAATATTAGTTG
Bam_5?gaq_fl   AGGCGAACTCCAGGAAGTC--GGCGAGGAGAAGCCGCGCTCAGCCATAC-----CG

Homom_gaq      ATCAAACAGCATGGCGTGCTGCTTAAGTGAAGAAGCCAAGGAACAGAAGAGGATAAACCA
Bam_5?gaq_fl   ACCTACGATGTTCACTTGCTGTTTGTCTCGGAGGAAGCTATACAGCAGATTAGAATCAATCA

Homom_gaq      AGAGATAGAACGACAATTAAGGAAGGATAAGAGAGATGCAAGAAAGAGAAGCTTAACTACT
Bam_5?gaq_fl   AGAAATTGAGAAGCAGCTAAAAAGAGACAAGAGAGATGCCCGAGAGAGCTTAACTGCT

Homom_gaq      CTTATTGGGCACTGGAGAGTCTGGCAAGTCAACATTCATCAACAGATGCGTATTATCCA
Bam_5?gaq_fl   TTTGTTGGGAACGG-----

Homom_gaq      TGGTGCTGGTTACAGTGATGAAGACAAGCGAGGTTTTATCAAGCTGGTCTT
Bam_5?gaq_fl   -----

```

B

```

      Start
Gaq_fl   MFTCCLSEEA IQQIRINQEIEKQLKRDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGV
Hom_Gaq  -MACCLSEEAKEQKRINQEIERQLRKDKRDARRELKLLLLGTGESGKSTFIKQMRIIHGA

Gaq_fl   GYTDDDRRDFTKLVYQNIFMAMQSMINAMELLKIEYKIPESKERAELVKAVDHETITRFE
Hom_Gaq  GYSDDEKRGFIKLVFNIFMAMQSMIRAMDLLQISYGDSANIEHADLVRSVDYESVTTFE

Gaq_fl   EPFVSAIRELWKDAGVQRCYERRNEYQLTDS TKYYLTDLERIA CRDYLPTQ QDILRVRAP
Hom_Gaq  EPYVTAMNSLWQDTGIQHICYDRRREYQLTDS AKYYLTDLDRIA AKDYVSTLQDILRVRAP

Gaq_fl   TTGIVEYPPFDLDS IFRMVDVGGQSRERRKWIHCFENVTSIIFLVALSEYDQILFESDSE
Hom_Gaq  TTGIIIEYPPFDLEEIRFRMVDVGGQSRERRKWIHCFENVTSIIFLVALSEYDQILFESDNE

Gaq_fl   NRMEESKALFR TIITYPWFQHSSVILFLNKKDLLEEKIMTSHLVDYFPPEYDGPQRDAIAA
Hom_Gaq  NRMEESKALFK TIITYPWFQHSSVILFLNKKDLLEEKIMYSHLVDYFPPEYDGPQRKDAIAA

Gaq_fl   REFILRMFVDLNPDA EKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV
Hom_Gaq  REFILRMFVELNPDEKIIYSHFTCATDTENIRFVFAAVKDTILQLNLKEYNLV

```

Fig. 10. Comparison between *B. improvisus* $G\alpha_q$ and *H. americanus* $G\alpha_q$ sequences. A, Comparison between 5' nucleotide sequences of $G\alpha_q$ is shown. B, comparison between $G\alpha_q$ amino acid sequences.

Amplification of $G\alpha_s$

To amplify the $G\alpha_s$ sequence, primers were designed based on ESTs from the *B. improvisus* cDNA library. In the $G\alpha_s$ EST sequences, parts of the 3' coding region is missing. The sequence

from the cDNA library has a longer open reading frame, than that in *H. americanus* $G\alpha_s$ with one ATG before the start ATG of *H. americanus* and the second ATG are positioned similar as in the *H. americanus* $G\alpha_s$. Two primers were designed, fw 1 before the first ATG and fw 2 before the second ATG. One rev primer located in the middle of the coding region, was designed and used in combination with two forward primers (Fig. 11).

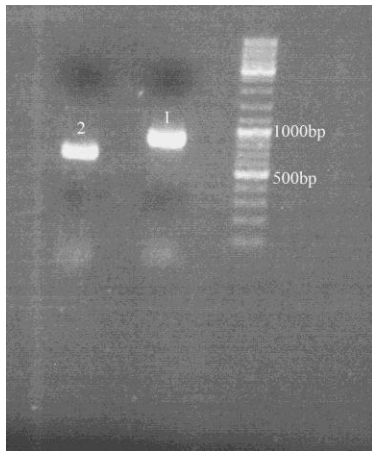


Fig. 11. PCR fragments of $G\alpha_s$. PCR was done by using two primer combinations fw 1, fw2 with rev primer. Two nice bands were obtained. PCR was run in the following way: 2 min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 62° C annealing for 30 s, 72° C extension for 1 min and final extension at 72° C for 10 min.

To amplify the 3' end of $G\alpha_s$, Gene racer 3' primer with nested 3' primer and GSPs (gene specific primers) race3' Gas1 and race3' Gas2 were used. After nested PCR three bands of about 1000 to 1500 bp were obtained (Fig. 12). The obtained products were purified from the gel, cloned and sent for sequencing. The 3' end of $G\alpha_s$ was amplified. Based on the sequence obtained from the middle band, reverse full length primer was designed to amplify the complete sequence of $G\alpha_s$. PCR was performed by using primers fw 1 and fw 2 with the rev full length primer, and the complete sequence of $G\alpha_s$ was amplified (Fig. 13). Three different variants of gas were found differing only in the 3' end of the coding region (Fig.14).

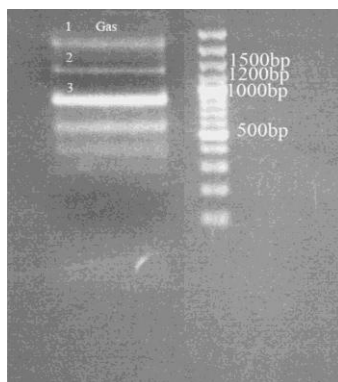


Fig. 12. 3' race for $G\alpha_s$. Three bands were obtained between 1000 to 1500 bp. The following PCR program was used: 2 min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 63° C annealing for 30 s, 72° C extension for 2 min and final extension at 72° C for 10 min.

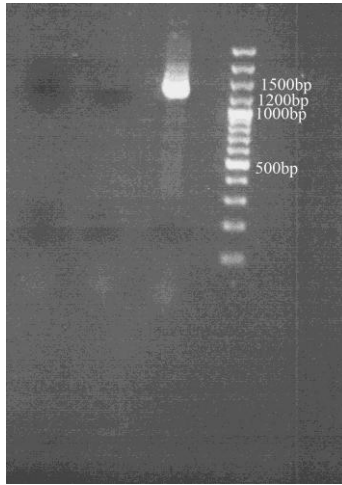


Fig. 13. $G\alpha_s$ PCR product obtained by using full length primers. Band was obtained between 1200 to 1500 bp. PCR program used was: 2 min denaturing step at 98° C followed by 35 cycles with a 98° C denaturing step for 30 s, 58° C annealing for 30 s, 72° C extension for 2 min and final extension at 72° C for 10 min.

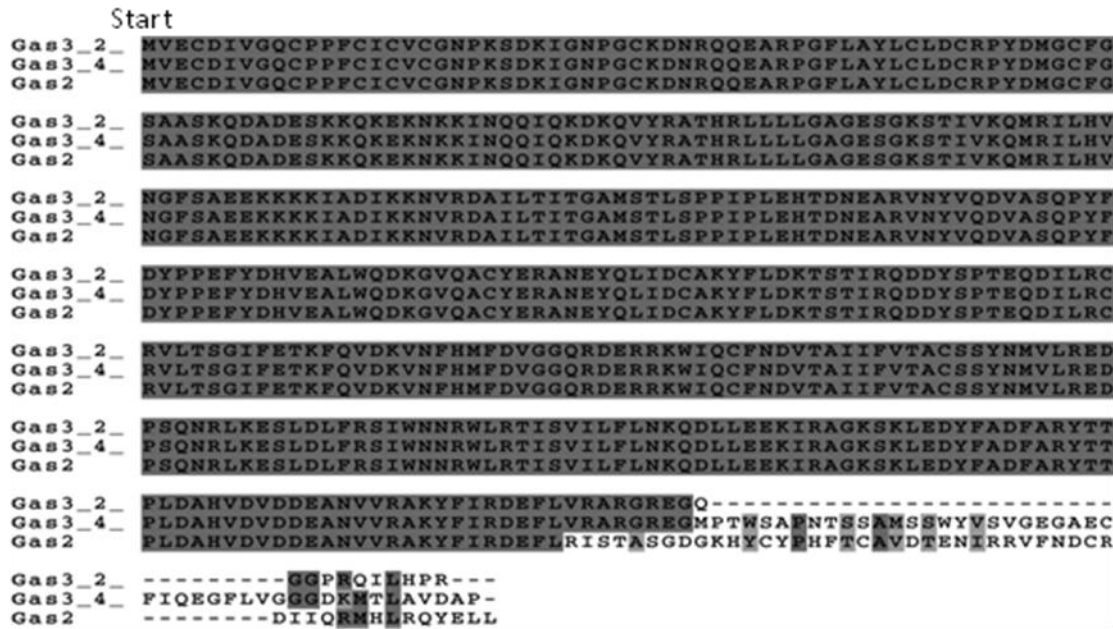


Fig. 14. An amino acid sequence analysis of 3 different clones of $G\alpha_s$ full length. The nucleotide sequences of three different clones of $G\alpha_s$ were translated and aligned together. After sequence alignment of $G\alpha_s$ different variants of $G\alpha_s$ were found. All of the three clones of $G\alpha_s$ show very high sequence similarity in the 5' end but in the 3' end upto 40 amino acid, they are different.

Sequence alignment of $G\alpha_s$ full length and *H. americanus*

Sequence alignment of $G\alpha_s$ full length and $G\alpha_s$ from *H. americanus* was done. It is found that both are relatively similar (Fig. 15).

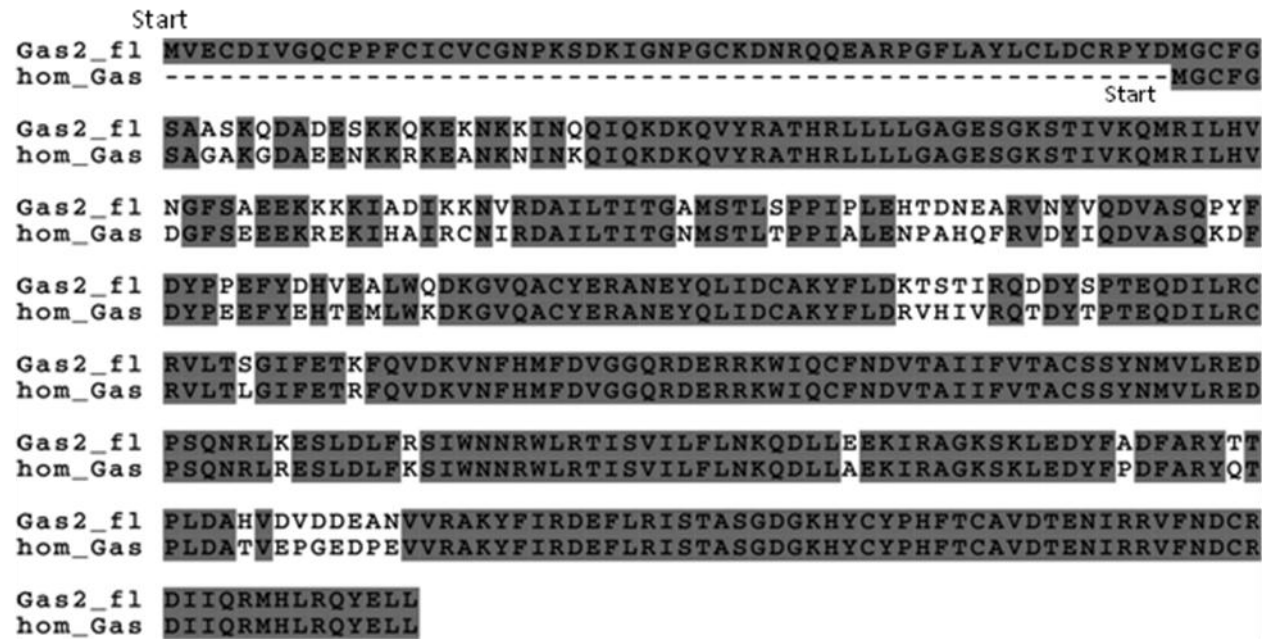


Fig. 15. Comparison between $G\alpha_s$ full length and *H. americanus* sequences.

Sequence identity of *B. improvisus* $G\alpha_q$ and $G\alpha_s$ with G-proteins of related organisms

The amino acid sequence of $G\alpha_q$ of *B. improvisus* shows high sequence similarity (83%) with the amino acid sequence of $G\alpha_q$ from *H. americanus*. The amino acid sequence of different variants of $G\alpha_q$ in Drosophila shows 65 to 83 % sequence similarity with $G\alpha_q$ of *B. improvisus*. The amino acid sequence of *B. improvisus* $G\alpha_q$ shows 68% sequence similarity with the amino acid sequence of Daphnia. The nucleotide sequence of one of the variants of $G\alpha_s$ (Gas2(6)) shows high sequence identity 80 to 86% with lobster, Drosophila and Daphnia. The other two variants of $G\alpha_s$ in *B. improvisus* show somewhat less sequence identity (Table 2).

Table 2. Identity table of $G\alpha_q$ and $G\alpha_s$

	GaS2(6)_Bimp	GaS3(2)_Bimp	GaS3(4)_Bimp	GaS_Dp	GNAS_HOMAM	GNAS_DROME	GAQ_49B_B_NM_165920_DROME	GAQ_49B_F_NM_165925.1_DROME	GaQ_Dp	Gaq_CG30054-PE_DROME	Gaq_CG30054-PF_DROME	Gaq_NP_001036539_CG17760_DROME
GaS2(6)_Bimp	100	96	90	84	86	80	43	43	38	39	39	37
GaS3(2)_Bimp	96	100	98	78	78	74	41	42	36	36	37	35
GaS3(4)_Bimp	90	98	100	73	72	69	39	40	34	34	35	34
GaQ_Bimp	43	42	39	68	83	42	83	76	68	67	66	65

Table 2. Amino acid identity table for $G\alpha_q$ and $G\alpha_s$ of *B. improvisus*, Drosophila, Daphnia and lobster. Bimp (*B. improvisus*), Dp (Daphnia), DROME (Drosophila) HOMAM (Homarus) and GNAS ($G\alpha_s$).

Phylogenetic analysis of *B.improvisus* $G\alpha_q$ and $G\alpha_s$

Comparison of $G\alpha_q$ and $G\alpha_s$ of *B. improvisus* was done by constructing a phylogenetic tree by using neighbor joining method. In order to construct a phylogenetic tree, nucleotide sequences of $G\alpha_q$ and $G\alpha_s$ obtained from other invertebrates like drosophila, daphnia and lobster were aligned with the barnacles nucleotide sequences. It is obvious from the tree that $G\alpha_q$ of *B. improvisus* has relationship with the $G\alpha_q$ of other organisms like daphnia, drosophila and lobster. The $G\alpha_s$ variants of *B. improvisus* also show close relationship with the $G\alpha_s$ of daphnia, drosophila and lobster.

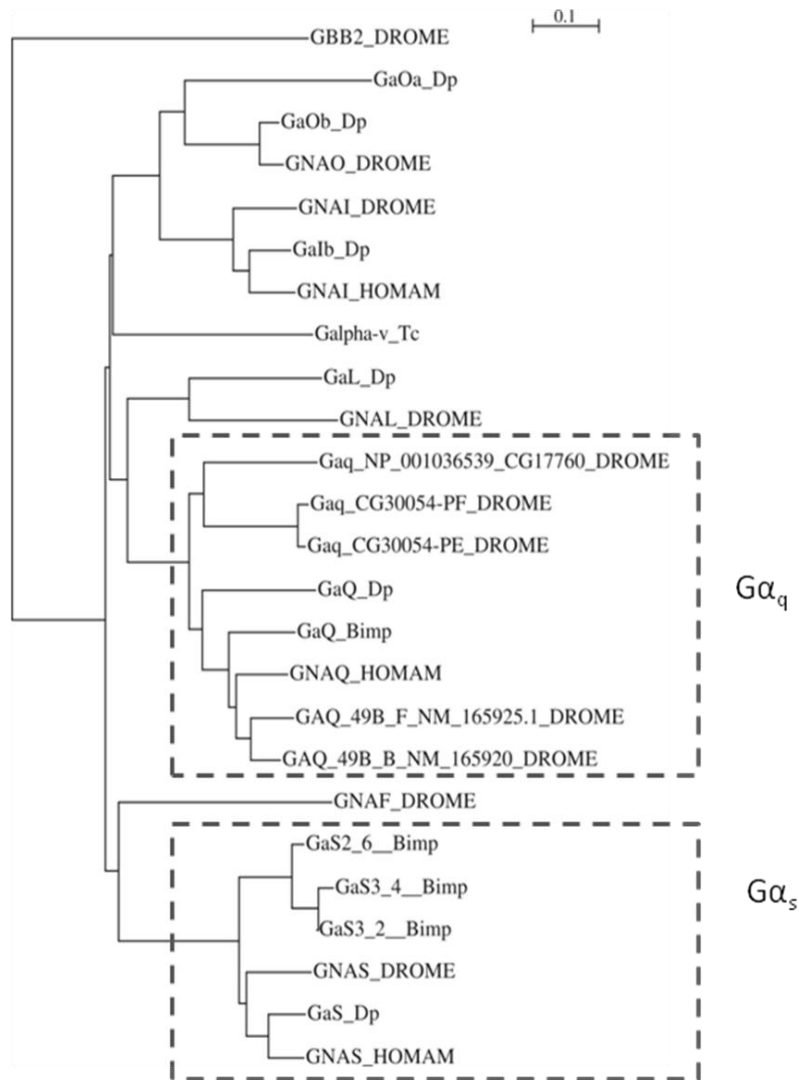


Fig. 15. A phylogenetic tree showing the relationship of *B. improvisus* $G\alpha_q$ and $G\alpha_s$ with the $G\alpha_q$ and $G\alpha_s$ of drosophila, daphnia and lobster. GBB2_DROME was used as an out group.

DISCUSSION

Barnacles, like many marine invertebrates cause serious biofouling problems. In this study the barnacle *B. improvisus* G-protein encoding genes $G\alpha_q$ and $G\alpha_s$ were cloned and sequenced. In a *B. improvisus* EST-library parts of the G-protein sequences were available. In order to amplify the full length proteins, PCR and RACE techniques were chosen to be the method of choice and cDNA was used as template to avoid introns. During the amplification of $G\alpha_q$ many difficulties were found, requiring design and optimization of several different primers and amplification conditions. The EST sequences covered the whole coding region except for a small gap in the middle of the gene, as judged by comparison with $G\alpha_q$ from *H. americanus*. Running PCR with primers based on EST sequences from just upstream of the predicted start codon and downstream of the stop codon, did not result in any product despite several optimization efforts. A possible explanation for this was obtained after performing 5' Race where the whole coding region and the complete 5' un-translated region was obtained. Comparing the 5' un-translated region of the EST sequences with the sequences from the RACE product showed that they were not similar. There is thus a possibility that there are two slightly different genes, or splice products of the $G\alpha_q$ in *B. improvisus*.

The two different 5' end sequences (5' race short and 5' race full length) of the $G\alpha_q$ of *B. improvisus* were aligned with *H. americanus*. The sequence obtained from the extreme 5' end of $G\alpha_q$ 5' race short was aligned with *H. americanus* sequence, but they were not very similar (about 50% identity). The nucleotides sequence alignment of the coding region of $G\alpha_q$ 5' race full length fragment with *Homarus. americanus* sequences, show that both sequences are relatively similar. It is concluded that $G\alpha_q$ has different variants of the 5' end, which is also seen in other $G\alpha$ subunits like $G\alpha_o$ (Hsu et al., 1990). One explanation could be that they are the products of different genes. The $G\alpha_s$ encoding gene of the fruit fly *Drosophila melanogaster* also has two variants of the N-terminus. In drosophila three genes encode $G\alpha_q$ and they also differ in N-terminus (www.flybase.org). The other explanation could be that these differences are due to variation found in the sequences from individuals of same species.

In the EST sequences of $G\alpha_s$ from *B. improvisus* the start ATG is present upstream as compared with the ATG of *H. americanus*, drosophila and daphnia. Whether the translation will start at the first ATG of the open reading frame in *B. improvisus* is not clear.

In case of $G\alpha_s$, three different variants were found which are very similar at the 5' end but differ largely in the coding region of the 3' end. The amino acid sequence of one of the variants is highly conserved and very similar with *H. americanus* sequences, while the two other variants are different in the C-terminal.

The occurrence of different variants of $G\alpha_s$ in *B. improvisus* at the 3' end might be due to alternative splicing at the 3' end of the mRNA. Another possible reason could be that these variants are the products of different genes. Splice variants of other G proteins also have been shown to have different amino acid sequences at the C-terminus. Three novel splice variants of $G\alpha_s$ were found in acute leukemia patients (Ye et al., 1999). The C-terminus of the G protein is known to interact with GPCRs in order to transmit signals (Yoo et al., 2002). We consider that change in the amino acid sequence at the C-terminus of $G\alpha_s$ would alter or abolish the interaction with the receptor, since these residues play a critical role in receptor interactions. It has been

shown that a $G\alpha_i$ protein C-terminal splice variant effect intracellular localization of dopamine receptor (Lopez-Aranda et al., 2007).

In the future, the other $G\alpha$ subunits of the G-protein like $G\alpha_i$ and $G\alpha_{12}$ will be cloned and sequenced. The interaction of different $G\alpha$ subunits with GPCRs will be studied by using the yeast two hybrid interaction system or other systems. The mRNA expression level of different G proteins in different tissues or during different developmental stages will be investigated by QPCR.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my Examiner Anders Blomberg for providing me an opportunity to work in his group, and for his valuable comments and suggestions on presentation and editing of thesis.

Sometimes words are not enough to express feelings, I am greatly indebted to my Supervisor Ulrika Lind for her patient guidance and generous assistance during lab work, for her help with the bibliography and diagrams.

I take this opportunity to thanks my sister for her patience and moral support.

Last but not the least I am very thankful to Michael, Magnus, Annabell and to all the group members for giving me a very good and cooperative working environment.

REFERENCES

- Callow, M. E., and Callow, J. E. (2002). Marine biofouling: a sticky problem. *Biologist (London)* **49**, 10-14.
- Han, D. W., and Watson, R. D. (2005). Trimeric G proteins in crustacean (*Callinectes sapidus*) Y-organs: occurrence and functional link to protein synthesis. *J Exp Zool A Comp Exp Biol* **303**, 441-447.
- Hsu, W. H., Rudolph, U., Sanford, J., Bertrand, P., Olate, J., Nelson, C., Moss, L. G., Boyd, A. E., Codina, J., and Birnbaumer, L. (1990). Molecular cloning of a novel splice variant of the alpha subunit of the mammalian Go protein. *J Biol Chem* **265**, 11220-11226.
- Isoai, A., Kawahara, H., Okazaki, Y., and Shizuri, Y. (1996). Molecular cloning of a new member of the putative G protein-coupled receptor gene from barnacle *Balanus amphitrite*. *Gene* **175**, 95-100.
- Kawahara, H., Isoai, A., and Shizuri, Y. (1997). Molecular cloning of a putative serotonin receptor gene from barnacle, *Balanus amphitrite*. *Gene* **184**, 245-250.
- Krang, A. S., and Dahlstrom, M. (2006). Effects of a candidate antifouling compound (medetomidine) on pheromone induced mate search in the amphipod *Corophium volutator*. *Mar Pollut Bull* **52**, 1776-1783.
- Lennquist, A., Hilvarsson, A., and Forlin, L. (2010). Responses in fish exposed to medetomidine, a new antifouling agent. *Mar Environ Res* **69 Suppl**, S43-45.
- Lind, U., Alm Rosenblad, M., Hasselberg Frank, L., Falkbring, S., Brive, L., Laurila, J. M., Pohjanoksa, K., Vuorenmaa, A., Kukkonen, J. P., Gunnarsson, L., *et al.* (2010). Octopamine receptors from the barnacle *balanus improvisus* are activated by the alpha2-adrenoceptor agonist medetomidine. *Mol Pharmacol* **78**, 237-248.
- Lodish, H., ed. (2008). *Molecular Cell Biology*, 6th edn.
- Lopez-Aranda, M. F., Acevedo, M. J., Gutierrez, A., Koulen, P., and Khan, Z. U. (2007). Role of a Galphai2 protein splice variant in the formation of an intracellular dopamine D2 receptor pool. *J Cell Sci* **120**, 2171-2178.
- Odling, K., Albertsson, C., Russell, J. T., and Martensson, L. G. (2006). An in vivo study of exocytosis of cement proteins from barnacle *Balanus improvisus* (D.) cyprid larva. *J Exp Biol* **209**, 956-964.
- Oldham, W. M., and Hamm, H. E. (2008). Heterotrimeric G protein activation by G-protein-coupled receptors. *Nat Rev Mol Cell Biol* **9**, 60-71.
- Tuteja, N. (2009). Signaling through G protein coupled receptors. *Plant Signal Behav* **4**, 942-947.
- Xu, F., Hollins, B., Landers, T. M., and McClintock, T. S. (1998). Molecular cloning of a lobster Gbeta subunit and Gbeta expression in olfactory receptor neuron dendrites and brain neuropil. *J Neurobiol* **36**, 525-536.
- Ye, Q., Lu, Q., Zhang, S., Huang, J., Wang, H., Su, G., Huang, P., and Huang, C. (1999). Existence of multiple novel Gs alpha splice variants in acute leukemia patients. *IUBMB Life* **48**, 299-304.

Yoo, J. H., Yang, Y. S., Choi, I., Shangguan, Y., Song, I., Neubig, R. R., and Wiley, J. W. (2002). Expression of novel splice variants of the G protein subunit, Go alpha, is tissue-specific and age-dependent in the rat. *Gene* 296, 249-255.