NEUROPEPTIDOME AND GPCROME OF STICK INSECT: SPECIFIC STRUCTURAL AND FUNCTIONAL ASPECTS OF ALLATOSTATIN RECEPTOR

by

Burçin Duan Şahbaz

B.S., Molecular Biology and Genetics, Boğaziçi University, 2009

M.S., Molecular Biology and Genetics, Boğaziçi University, 2012

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Molecular Biology and Genetics

Boğaziçi University

2018

To my family and my sister...

Just stay where you are.

ACKNOWLEDGEMENTS

The data in this project were produced in collaboration with the BioAFM Laboratory of Hamdi Torun in Boğaziçi University, SezermanLab of Uğur Sezerman in Acıbadem University and Vivarium Facility of Boğaziçi University. All of the molecular biology techniques were performed in Cancer Signaling Laboratory (CSL) of Necla Birgül İyison. Atomic Force Microscopy measurements were conducted in BioAFM Laboratory. Assembled transcriptome was generated in and bioinformatic predictions were consulted from SezermanLab. The cancer cell lines which were utilized in viability assays were the gifts of Nesrin Özören and Huh7 cell line was of Mehmet Öztürk. I would like to thank the Vivarium team for all their help, even at nights, and Hamdi Torun for opening his lab and giving his valuable time. I would like to extend my special thanks to Uğur Sezerman, who supervised my bioinformatics work for almost six years and who did not hesitate to give me his precious time, supported and encouraged me all the time. The studies in this thesis were founded by Boğaziçi University Scientific Research Projects (project codes: 14B01P2 and 17B01D1) and Tübitak (project code: 113Z079). I would also like to thank the other members of the thesis committee, İbrahim Yaman, Serdar Durdağı and Igor Kryvoruchko. I am especially grateful to Ibrahim Yaman for making the years fun and pleasant. Every time I was in trouble, he was one of the people with an open door.

I feel very grateful and lucky to become a member of the CSL family and to meet and share anything with the former and recent members of the lab. They all gave me a huge world to discuss, share, criticize and even struggle with. I learnt too much from Tuncay Şeker, shared grate time with İzzet Akiva, Aida Shahraki and the others. Also some unofficial members such as Güneş Tunçgenç and the other graduate students have added their mind, energy and data.

I would like to point out that none of these studies could have emerged if some people were not with me. These people are my mother who was beside me anytime I asked, my husband whom I love too much and who never gave up on me even in the

most stressful days, my sister who is the ghost assistant of these people, my beloved daughter who distracted me and the other members of my family. I owe them all those years.

Lastly, I thank Necla Birgül İyison whose contribution is invaluable to my life, carrier and also to this thesis. She opened the door for the first time and gave me the freedom to imagine, design and work. This opportunity helped me to become really experienced on what I've done. She was always kind, encouraging, supportive, sometimes as close as a friend sometimes as a mother, kind-hearted and a very good cook.

Thanks for any minute that I spent in CSL.

ABSTRACT

NEUROPEPTIDOME AND GPCROME OF STICK INSECT: SPECIFIC STRUCTURAL AND FUNCTIONAL ASPECTS OF ALLATOSTATIN RECEPTOR

The arthropods constitute about three-quarters of the world's animal species. The molecules that regulate many physiological events, from the feeding to development, from locomotion to the social behavior, from the reproductive behavior to the intestinal motility, are the neuropeptides and their cognate receptors. *Carausius morosus*, also known as laboratory rodent, is a species that is studied on locomotion and can easily reproduce via parthenogenesis. On the other hand, among the arthropod neuropeptides, those that regulate Juvenile Hormone, namely allatostatin (AST), is specially important. The subject of this thesis is understanding the interaction between allatostatin C receptor (AlstR-C) of *C. morosus* and AST-C, as well as finding the other neuropeptides and GPCRs. At the beginning of our work, it was found that amino acids were conserved in the ligand binding pocket of AlstR-C and these amino acids were mutated and utilized in atomic force microscopy studies. This IXTPP motif, located in the third extracellular loop, together with the N-terminus were found to be important for this interaction. RNA sequencing analysis was then performed to access other AlstR types and AST peptides. As a result,at least 23 different neuropeptide transcripts and 43 GPCR transcripts were obtained from adult *C. morosus*. Tissue expression profiles of these GPCRs were found. This information will facilitate future neuropeptide-GPCR studies. In this study, about the Alstr-AST system being homologous to the somatostatin receptor of human, we have also asked whether this neuropeptide may affect the proliferation of cancer cells. However, both XTT and *in vivo* xenograft experiments showed that the peptide or active receptor does not affect tumor growth.

ÖZET

ÇUBUK BÖCEĞİNİN NÖROPEPTİDOM VE GPCROM ÇALIŞMASI: ALLATOSTATİN RESEPTÖRÜ ÖZELİNDE YAPISAL VE İŞLEVSEL ÇALIŞMALAR

Eklembacaklılar, dünyadaki hayvan türlerinin yaklaşık dörtte üçünü oluşturur. Bu organizmaların embriyolojik gelişimden beslenmesine, hareket yeteneğinden sosyal davranışlarına, bağırsak hareketlerinden üreme davranışlarına kadar pek çok fizyolojik olayını düzenleyen moleküller nöropeptidler ve bunların ilgili reseptörleridir. Laboratuvar çubuk böceği olarak bilinen *Carausius morosus* ise hareket yeteneği üzerine çalışılan ve partenogenetik özelliğiyle çok kolay üreyebilen bir türdür. Öte yandan eklembacaklı nöropeptitleri arasında Juvenil Hormonu regüle edenlerin yani allatostatinlerin (AST) özel bir önemi vardır. Bu tezin konusu da *C. morosus*'un hem allatostatin C reseptörü (AlstR-C) ve AST-C bağlanmasının anlaşılması hem de diğer nöropeptit ve GPCRlerin ortaya çıkarılmasıdır. Çalışmamızın başlarında AlstR-C'nin ligandı ile bağlanma cebinde korunmuş amino asitler olduğu görülmüş ve bu amino asitler mutasyona uğratılarak atomik kuvvet mikroskobu çalışmalarında kullanılmıştır. Dış çevrimlerden üçüncüsünde yer alan bu IXTPP motifi ve N-ucunun ligand bağlanmasında önemli olduğu yapılan deneylerde görülmüştür. Daha sonra diğer AlstR tiplerine ve AST peptitlerine ulaşabilmek için RNA dizileme analizi yapılmıştır. Bunun sonucunda yetişkin *C. morosus*'ta anlatılan en az 23 farklı nöropeptit transkripti ve 43 tane de GPCR transkripti elde edilmiştir. Bu GPCRlerin hangi dokularda anlatıldığı bulunmuştur. Bu bilgiler gelecekteki nöropeptid-GPCR çalışmalarına yardımcı olacaktır. Bu çalışmada AlstR-AST sisteminin insandaki somatostatin reseptörü ile benzerliğinden yola çıkılmış ve bu kanser hücrelerinin çoğalmasını etkileyip etkilemeyeceği de sorulmuştur. Fakat elde edilen sonuçlara göre ne aktif reseptör ne de peptidin kendisi tümör gelişimini durdurmamaktadır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

g	gram
l	liter
m	meter
n	nano
p	pico
U	unit
α	alpha
$\mathring{A}$	Ansgtrom
β	beta
δ	delta
μ	micro
®	registered trademark

LIST OF ACRONYMS/ABBREVIATIONS

3D	Three Dimensional
7TM	7-Transmembrane
AFM	Atomic Force Microscopy
AlstR	Allatostatin Receptor
AST	Allatostatin Peptide
BPB	Bromophenol Blue
BSA	Bovine Serum Albumin
Ca	Calcium
CA	Corpus Allatum
cAMP	Cyclic Adenosine Monophosphate
CC	Corpus Cardiacum
cDNA	Complementary Deoxyribonucleic Acid
CDS	Coding Sequence
C-term	C-terminal Loop
DAPI	4′,6-diamidino-2-phenylindole
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide
ECL	Extracellular Loop
EDTA	Ethylenediaminetetraacetic Acid
ERK	Extracellular Signal Regulated Kinase
EtBr	Ethidium Bromide
EtOH	Ethanol
FBS	Fetal Bovine Serum
FRET	Förster Resonance Energy Transfer
GFP	Green Florescent Protein
GPCR	G Protein-Coupled Receptor

HA	Hemagglutinin
hr	hours
ICL	Intracellular Loop
IF	Immunofluorescence
JH	Juvenile Hormone
Koff	Dissociation Rate Constant
LB	Luria-Bertani Broth
MD	Molecular Dynamics
min	minutes
MOPS	3-(N-morpholino) Propanesulfonic Acid
mRNA	Messenger Ribonucleic Acid
NaOAc	Sodium Acetate
Ndel	N-terminus Deletion
Nterm	N-terminal Loop
O/N	Overnight
ORF	Open Reading Frame
PCR	Polymerase Chain Reaction
PI	Propidium Iodide
RMSD	Root Mean Square Deviation
rpm	Revolutions per Minute
qPCR	Quantitative PCR
RT	Room Temperature
RT-PCR	Reverse Transcription PCR
SDS	Sodium Dodecyl Sulphate
sec	Seconds
SRA	Sequence Read Archives
SSTR	Somatostatin Receptor
TEMED	Tetramethylethylenediamine
TGF-α	Transformin Growth Factor Alpha
TM	Trans-membrane
XTT	2,3-bis-(2-Methoxy-4-Nitro-5-Sulphenyl)-(2H)-Tetrazolium-5-Carboxanilide

WT	Wild Type
ZO1	Zonula Occludens-1

1. INTRODUCTION

1.1. Neuropeptides and G Protein-Coupled Receptors

Neuropeptides are the peptide hormones that are produced, processed in and secreted from the neurons. Majority of the neuropeptides exert their physiological functions through binding to specific G protein-coupled receptors (GPCR). These functions range from development to behavior. For instance, they can simply work for the tentacle contraction of the *Hydra* which does not have a brain, post-natal baby care of some arthropod species and also dropping the tail of the lizard. The hypothesis on the diversity of neuropeptides depends on tandem duplications and substitutions. Therefore, small numbers of genes code for a large neuropeptide repertoire [1].

On the other hand, peptide GPCRs are members of the largest two GPCR families, namely Class A and Class B1. Class A GPCRs are called as "rhodopsin-like" while Class B1 as "secretin-like". In case of insects, the neuropeptides such as allatostatin (AST), adipokinetic hormone (AKH), myoinhibitory peptide (MIP) and most of the others couple with Class A GPCRs, while a few other neuropeptides such as calcitonin and diuretic hormone (DH) couple with Class B1 GPCRs. Binding of the neuropeptide to its cognate GPCR leads to a conformational change on the receptor which in turn activates various intracellular signal transduction pathways. This transduction is generally defined by the type of Gα protein that is recruited. However, the fundamentals of pathway preference is much more complicated than this.

For many years, insect neuropeptides and their GPCRs have helped to model mammalian endocrine systems, due to the functional and sequence-based similarity within species. However, these systems are different in other aspects, on the other hand. For example, there are neuropeptides that take part in specific mechanisms only for insects. In addition, many neuropeptides active in insects have homologous GPCRs in mammals but cannot activate them. This allows neuropeptides to be safely used in pest control studies.

The walking stick insect *Carausius morosus* is a laboratory insect that belongs to the Phasmatodea order. It is herbivorous and has important roles in the ecosystem. On the other hand, its outbreak damages the park areas or plantations of economic importance. In biological research, it is commonly studied for its specific locomotion behavior [2] and its close relatives are for their camouflage ability. Recently, the neuropeptides responsible for this specific locomotion behavior began to attract attention [3].

1.2. Allatostatins, Their Receptors and The Structure of C-type

Ecdysones and Juvenile Hormone (JH) are the key regulators of insect developmental stages. Within these, JH acts in a broad range of mechanisms such as reproduction, feeding, aggressiveness, post-natal care, oogenesis, vitellogenesis etc, and its synthesis is tightly regulated by ASTs secreted from the neurosecretory cells extending from brain to the neuroendocrine gland *corpus allatum* (CA).

There are three types of ASTs. These are represented as A, B and C types. The ligands of each type have conserved C-terminal core regions. A types have a common FGL-amide pattern, B types W(X)6W-amide while C types PISCF-amide on the C-terminus (C-term). Each type of peptide stimulates a corresponding receptor. Allatostatin receptors (AlstR) are the cognate GPCRs of this peptide class. They were firstly identified from fruit fly in the late 1990s [4] by using their evolutionary homology to somatostatin receptors (SSTR) of mammals. They have other roles, such as inhibition of muscle contraction in the midgut, regulation of the feeding behavior and the reproductive ability of the insects [5,6], directly or via regulating JH. The type of receptor and its related function vary among species.

Until 2017, the information on the direct interaction and kinetics of AlstR and ASTs was very limited. The first structural and pharmacological studies on the C type of AlstR were performed in our laboratory [7, 8]. We have discovered AlstR-C in *C. morosus* and found its ligand binding pocket. *In silico* analyses showed a conserved putative binding motif (Figure 1.1) in the extracellular loop 3 (ECL3). Figure 1.2c

shows the Hydrogen bonds (blue lines) and Van-der Waals forces (dotted yellow lines) between the residues that are responsible for the interaction. These residues constitute the N-terminal loop (N-term) and ECL2 (Figure 1.2d and e) as the other parts of the binding pocket.

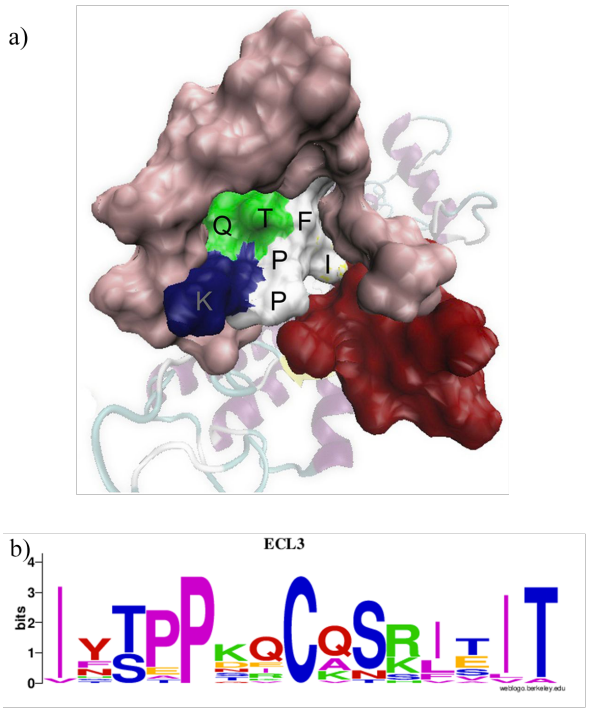


Figure 1.1: The localization and conservation of IXTPP motif. (a) 3D representation of IFTPP residues on the ECL3. ECL3 was surrounded by ECL2 (red) and N-term (pink). (b) Sequence logo of the amino acids on ECL3, processed in weblogo tool of Berkeley University. The figure was adapted from [7].

The following studies in our laboratory revealed that CamAlstR-C could couple with Gα_{i1}, decrease cAMP accumulation and recruit β-arrestins upon activation with AST-C [8]. In addition, the mutations on IFTPP residues of ECL3 lead to a decrease in activity and in the recruitment of β-arrestin.

All of this has led us to learn more about and to support the importance of this IXTPP motif.

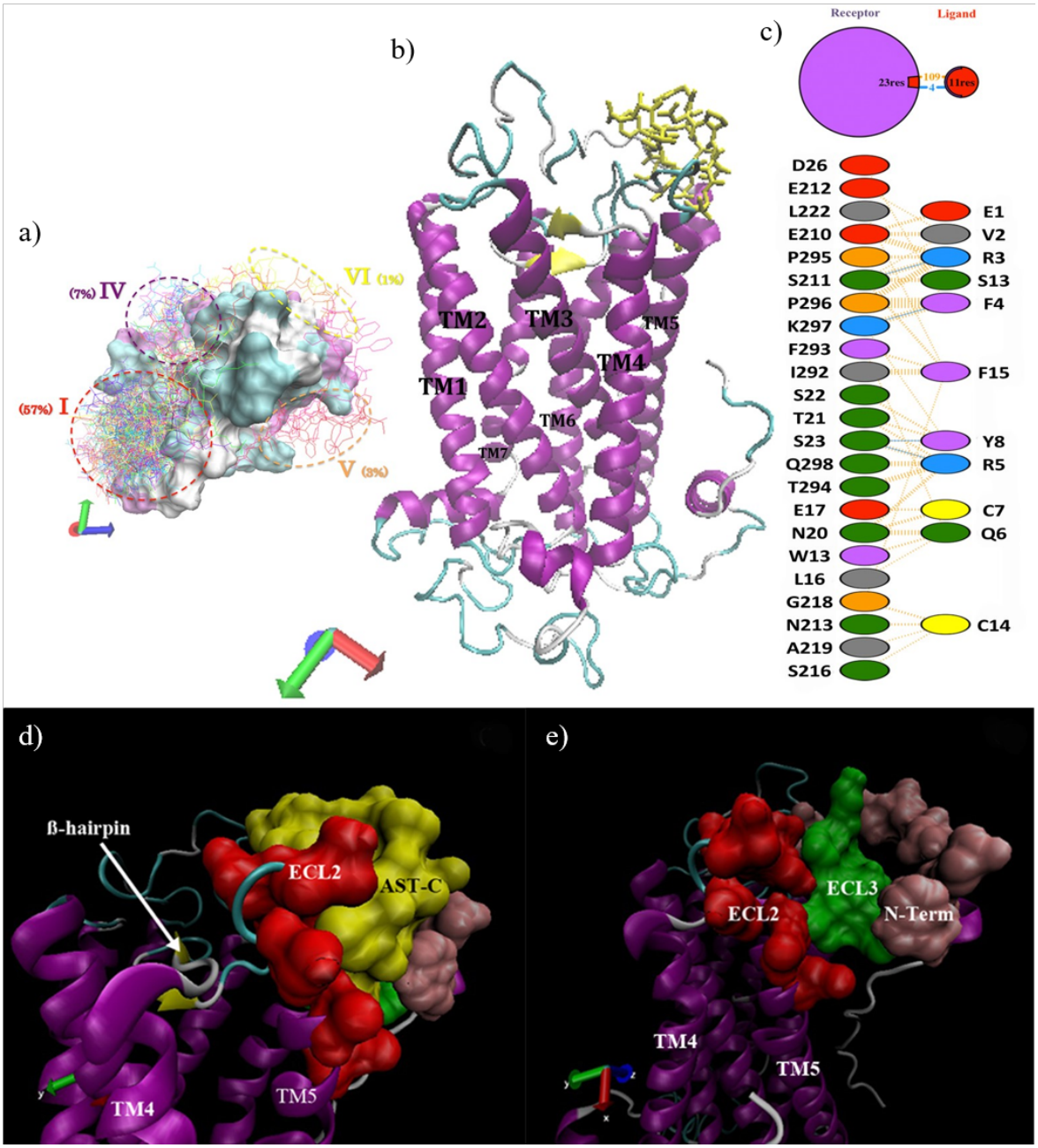


Figure 1.2: Docking results of CamAlstR-C with AST-C. (a) Clustering analysis of different docking poses (n=100). TM: transmembrane helix. (b) 3D structure of the complex. (c) The residues responsible for the interaction. The 3D representation of binding pocket (d) with and (e) without the ligand. The figure was adapted from [7].

1.3. Allatostatin Receptor Type C as The Homolog of Human Somatostatin Receptor

The members of AlstR family are functionally and structurally homologous to different mammalian GPCRs. For instance, A type is closely related to galanin, B type to bombesin and C type to somatostatin receptor. These mammalian receptors have distinct functions. Galanin receptors are known to inhibit action potentials in neurons and function in feeding, regulation of mood and sleep/awakeness [9]. In correlation with this function, the first AlstR-A of *D. melanogaster* was utilized as a neuronal silencer in mice [10] and macaque monkey [11]. Bombesin receptors block eating behavior [12] and its homolog AlstR-B has myoinhibitory roles on gland and muscle activities [13]. SSTRs regulate growth hormone and leptin release from hypothalamus [14], AlstR-C inhibits JH and regulates feeding behavior of insects.

AlstRs show $\geq$70% similarity on their transmembrane (TM) region with SSTRs which belong to somatostatin/opioid subfamily of Class A GPCRs. There are five subtypes of human SSTRs that contribute to a sequence pattern in their 7-transmembrane (7TM) region, YANSCANPI/VLY. They have the common DRY sequence on their third trans-membrane helix (TM3) as with the other Class A GPCRs. The second and third intracellular loops (ICL2, ICL3) interact with G proteins. Ligand binding region consists of ECL2, ECL3, and parts of TM6 and TM7 [15]. According to the information obtained from agonist studies, different regions of SSTRs are important for ligand-specific interaction.

Both SST and AST-C peptides include a disulfide-bridge to become biologically active. SST peptides have a conserved FWKT motif, but AST-C peptides have PISCF motif on their C-term. Both inhibit the secretion of specific hormones that are responsible for the growth of the animal. SST inhibits the growth hormone and AST-C inhibits the JH.

SSTRs are expressed in various tissues such as the central nervous system, pancreas, stomach, and intestines. The first SST (SST-14) was named after somatotropin-

release due to its inhibiting role of the hypothalamic growth hormone release [16]. This action was not the only effect of SSTR-SST activity. According to the following studies on these receptors and peptides since the 1970s, SSTs result in various inhibitory responses such as inhibition of hormone secretion or proliferation [17–19]. This inhibitory role made SSTRs valuable targets for agonist treatment in cancer cases [20–22]. SST analogs are used against neuroendocrine tumors (NET) such as gonadotropic adenomas and pancreatic tumors. Octreotide is one of the most known SST analogs. Lanreotide is the secondly used agonist for SSTRs [23] and pasireotide is the other which is still being tested in a clinical trial [24]. The data on trials of these analogs exhibited tumor shrinkage, even disappearance. Some trials showed tumor stabilization. Their macro effects are proposed as inhibition of growth factors [25] and angiogenesis [26]. Additionally, cellular effects are proposed as cell cycle arrest [17] and pro-apoptotic responses [27].

On the other hand, neither AlstR-C nor its ligand AST-C is expressed in mammalian cells. However, their mechanism of action resembles SSTR and SSTs. Insect AlstR-C inhibits JH secretion, reduce muscle contractions in the midgut and inhibit vitellogenesis [6]. Therefore it became a valuable target for pesticide design [28]. These all showed that AlstR-C could exhibit similar functions to that of SSTR. Here we propose an inhibitory mechanism for AlstRC-ASTC activity against cell proliferation in mammalian cancer cells. We have previously shown that AlstR-C can couple with mammalian G proteins, especially Gα_{i1}, in 293FT cell line [8]. Additionally, we found that neither AST-C peptides could activate SSTR and nor SST ligands activate AlstR-C [29] or AST-C could not interact with any other surface proteins on Huh7 cell lines [7]. Therefore, ectopic expression of AlstR-C in mammalian cancer cells can make these cells specific targets for AST-C treatment, which might finally inhibit tumor growth *in vivo*.

2. PURPOSE

The actions of AlstR-C are associated with the actions of JH and this hormone is becoming increasingly popular day by day. JH analogs have economic importance in the market. Although AST is the only known inhibitor of JH synthesis, a link has not been established in between the actions of JH and AlstRC-ASTC mechanism. On the other hand, SST analogs are used against NETs. Therefore, we predict that AlstR-C and ASTC would take their place in the near future, in terms of both a system and an agonist design. This prediction led us to study the interaction between AlstR-C and its ligand, ASTC. Together with the previous studies performed in our laboratory, we aimed to find its binding pocket, the own ASTC peptide, the other types of AlstRs, their expression profiles and proposed a possible inhibitory function to this receptor-ligand system in cellular level via its homology to somatostatin receptor.

3. MATERIALS

3.1. Chemicals

The chemicals used in this study are the molecular biology grade chemicals purchased from VWR (PA, USA), AppliChem (USA), Sigma Aldrich (MO, USA), BioLine (Toronto, Canada) and Thermo-Fisher Scientific (MA, USA).

3.2. Kits

For polymerase chain reaction (PCR), DNA and RNA isolation and cell viability commercially available kits were used. For the amplification of specific mutant sequences and the coding sequences that will be used in transfection experiments, high fidelity enzymes are used such as Phusion High Fideliy PCR Kit (E0553S, New England Biolabs, MA, USA). For the other random PCRs, Taq DNA Polymerase With Standard Buffer (M0273S, New England Biolabs, MA, USA) and Taq DNA Polymerase, recombinant (EP0402, ThermoFisher Scientific , MA, USA) were used. For plasmid purification, NucleoSpin Plasmid (740588, Macherey-Nagel, UK) and for RNA isolation RNeasy Mini Kit (74160, QIAGEN, Hilden, Germany) were used. For cDNA synthesis ImProm-II Reverse Transcription System from Promega (Madison, WI, USA) and SensiFAST cDNA Synthesis Kit (BIO-65054, Toronto, Canada) were used. For qPCR, SensiFAST SYBR No*ROX Kit (BIO-98050, Bioline, Toronto, Canada) was used. For cell viability assays, Cell Proliferation Kit II (XTT) from Sigma-Aldrich (11465015001, MO, USA) was used.

3.3. Enzymes

For molecular cloning purposes, T4 DNA Ligase (M0202S, New England Biolabs, MA, USA) was used. And the restriction enzymes were bought as FastDigest enzymes of Thermo-Fisher Scientific (MA, USA). Trypsin (0.025 per cent, ready to use) was purchased from Gibco (Paisley, UK). pLENTI-III-HA vector system was purchased

from ABM, Inc. (BC, USA). mCherry2-N1 was obtained from Addgene. pcDNA3-SYFP2 was previously prepared in our laboratory. RNase A (EN0531) was purchased from Thermo-Fisher Scientific (MA, USA).

3.4. Nucleic Acids

DNA ladders used in the study are Gene-On DNA Markers 100 bp (304-005) and 1 kb (305-005). dNTP mix (R0191) were purchased from Thermo-Fisher Scientific (MA, USA).

3.4.1. Plasmids

pEGFP-N2 (Clontech, CA, USA), pcDNA3 (Invitrogen, CA, USA) plasmids were commercially obtained. pcDNA3/HA plasmid was obtained from the laboratory of Nesrin Özören, Molecular Biology and Genetics, Boğaziçi University.

3.4.2. Primers

Primers used in polymerase chain reactions, sequencing and cloning were purchased from Macrogen Europe (AZ, Netherlands). Primer sequences and codes used in the site-directed mutagenesis experiments were given in Table 3.1. The nucleotides in Bold correspond to the mutated nucleotides. The primers that were used in expression analyses were given in Table 3.2.

Table 3.1: List of primers used in cloning and site-directed mutagenesis.

Mutation code	**Primer code**	**Sequence (5'-3')**
WT	WT-ATG	AAAGCTTATCTAGAAAAATGTCTGTGGAACAAGTGACG
WT	WT-stop	TTTGAATTCTTGGATCCTCTACACCTGGGTCGGCTG
Ndel	Ndel-ATG	AAAAAGCTTATGGACACAGACCAGCCGACG
AFTPP	F1	ACGCAGATGGCGCTC**GC**CTTCACGCCGCCCAA
AFTPP	R1	TTGGGCGGCGTGAAGGCGAGCGCCATCTGCGT
AFTPA	F11	TCATCTTCACGCCG**G**CCAAGCAGTGCCAGT
AFTPA	R11	ACTGGCACTGCTTGGCCGGCGTGAAGATGA
AATPA	F111	AGATGGCGCTCGCC**GC**CACGCCGGCCAA
AATPA	R111	TTGGCCGGCGTGGCGGCGAGCGCCATCT
AFAPA	F112	ATGGCGCTCGCCTTC**GCC**CCGGCCAAGCAGT
AFAPA	R112	ACTGCTTGGCCGGGGCGAAGGCGAGCGCCAT
AFAAA	F1121	TCGCCTTCGCC**GCC**GCCAAGCAGT
AFAAA	R1121	ACTGCTTGGCGGCGGCGAAGGCGA
AAAAA	F11211	AGATGGCGCTCGCC**GC**CGCCGCCGCCAAGC
AAAAA	R11211	GCTTGGCGGCGGCGGCGGCGAGCGCCATCT

The bold characters represent the site of mutation.

Table 3.2: List of primers used in expression analysis of GPCR transcripts.

Primer Code	**Sequence (5'-3')**	**Name of GPCR**
CamActin-forw	AACTTCCTGATGGCCAGGTC	
CamActin-rev	ATGTCCACGTCGCACTTCAT	
CamG3PDH-forw	ACGGCGTCGAAGCAAAATTC	
CamG3PDH-rev	CGGCAGGTACTTGACGTTCT	
CamGAPDH-forw	CACTAAAGGGCATCCTGGCA	
CamGAPDH-rev	ATGGCATTGGGAGGAGAAGC	
CamAlstRC-forw	GCCTTCACGCTCTACACCTT	Allatostatin C Receptor
CamAlstRC-rev	GTGATGGACACCCTCGACTG	Allatostatin C Receptor
19522_c0_g1_i1-forw	CAGCAAGGGCCACCAGATAG	Adhesion GPCR G2-like
19522_c0_g1_i1-rev	GACGCTAGCAGGGAGTAGTG	Adhesion GPCR G2-like
21880_c0_g1_i1-forw	CGACATGGTGTTCGCCCT	Neuropeptide Y Receptor
21880_c0_g1_i1-rev	GACGGGGTGTTAGACAGGAG	Neuropeptide Y Receptor
28926_c1_g2_i1-forw	GCTGGCCTTCCTTGCACTAT	Frizzled 10
28926_c1_g2_i1-rev	CACGTACAACGCAGCAAACA	Frizzled 10
29760_c0_g1_i1-forw	AAGGCGCCCATAATCTTCGT	Diuretic Hormone Receptor
29760_c0_g1_i1-rev	CTCCGTGTTCAGGAAGCAGT	Diuretic Hormone Receptor
30951_c0_g1_i2-forw	TGATGGGTTCTCATGTGCCC	Octopamine Receptor
30951_c0_g1_i2-rev	GCCATTCACAATGCCCATCC	Octopamine Receptor

Table 3.2: List of primers used in expression analysis of GPCR transcripts (cont.)

Primer Code	**Sequence (5'-3')**	**Name of GPCR**
31442_c1_g1_i3-forw	GCATCATCATGGGCGTGTTC	Tyramine Receptor 2
31442_c1_g1_i3-rev	TTGGCCTGCGCAAGGTATTA	Tyramine Receptor 2
34134_c0_g1_i1-forw	TCTCGTACCAGGAAACTGCG	Gustatory Receptor for Sugar Taste 43a-like
34134_c0_g1_i1-rev	ACAGCACCGACAAATACGGT	Gustatory Receptor for Sugar Taste 43a-like
34460_c0_g1_i1-forw	CTTCATCGCGTCGCTCTACT	Neuropeptide CCHamide receptor
34460_c0_g1_i1-rev	CCATGTGGGCAGAGTTCCTC	Neuropeptide CCHamide receptor
35009_c0_g2_i1-forw	CAAGGGAAAACATCGCCTGC	Cholecystokinin receptor
35009_c0_g2_i1-rev	TGAACGAGTAGTGCCCGAAC	Cholecystokinin receptor
35728_c0_g1_i3-forw	CTGTGGTACCGCCTCATTGT	Calcitonin Gene-Related Peptide Type 1 Receptor
35728_c0_g1_i3-rev	AGGAAGCCGAGGTTTAGCAC	Calcitonin Gene-Related Peptide Type 1 Receptor
36849_c0_g1_i5-forw	TCTCGTACCAGGAAACTGCG	Inotocin Receptor
36849_c0_g1_i5-rev	GATCTTGGCTCGGGAGATGG	Inotocin Receptor
36998_c1_g1_i1-forw	TTGAAGGGACAGAACGCCAG	Adipokinetic Hormone Receptor
36998_c1_g1_i1-rev	TTCTGTACCACCGGGTTCAC	Adipokinetic Hormone Receptor

Table 3.2: List of primers used in expression analysis of GPCR transcripts (cont.)

Primer Code	**Sequence (5'-3')**	**Name of GPCR**
54154_c0_g1_i1-forw	ACGCGGAACGAGAAGAAGAA	Sex Peptide Receptor
54154_c0_g1_i1-rev	AGTACGTTCCGCTGGACATC	Sex Peptide Receptor
62595_c0_g1_i1-forw2	GTGCAGTACCTGATCGTCGT	Allatostatin A Receptor
62595_c0_g1_i1-rev2	ATGTAGTTGGTGACGCCGTG	Allatostatin A Receptor
65134_c0_g1_i1-forw	TGCGTTTACAACTGGTGGGA	Orphan GPCR
65134_c0_g1_i1-rev	AATGAGCCTTGCCTCGTTGT	Orphan GPCR

3.5. Peptides and Antibodies

Synthesis of AST-C of *D. melanogaster* was ordered from Biomatik (Ontario, Canada) and of *C. morosus* from SynPeptide Co. Ltd. (Shanghai, China). The antibodies used in the study are ZO-1 antibody (40-2200, Invitrogen, MA, USA), HA-Tag (6E2) Mouse mAb (2367, Cell Signaling, MA, USA), Anti-rabbit IgG, HRP-linked Antibody (7074, Cell Signaling, MA, USA), Anti-mouse IgG, HRP-linked Antibody (7076, Cell Signaling, MA, USA) and Goat Anti-Rabbit IgG H&L (Alexa Fluor® 555) (ab150078, abcam, Cambridge, UK).

3.6. Bioinformatics Tools

Trinity, TopHat, tblastx, tblastn, ORFPREDICTOR, TMHMM, SignalP 4.1, WorkBench, ExPASy, SMART, GPCRPred were utilized in bioinformatics for various procedures.

3.7. Bacterial Strains

Bacterial strain used in this study was *E. coli* DH5α (genotype: F-Ψ80d lacZΔM15-Δ(lacZYA-argF) U169 end A1 recA1 hsdR17 (rk-, mk+) supE44Λ- thi-1 gyrA96relA1 phoA).

3.8. Cell Lines

Huh7 was kindly provided by Prof. Dr. Mehmet Öztürk and MeWo, MCF7, HeLa and HepG2 cell lines were kindly provided by Prof. Dr. Nesrin Özören. 293FT cell line was previously grown in our laboratory.

3.9. Buffers, Media and Recipes

3.9.1. DNA Gel Electrophoresis

50X Tris-acetic acid EDTA (TAE)	2M Tris-acetate 50mM EDTA pH 8.5
EtBr	10 mg/ml
Loading buffer	For 10ml: 2.4 ml dH2O 0.1 ml 1M Tris-HCl, pH 7.6 0.3 ml 1 per cent Bromophenol Blue (BPB) 6 ml 100 per cent glycerol 1.2 ml 0.5M EDTA

3.9.2. RNA Gel Electrophoresis Buffers

DEPC treated	1 per cent (v/v) DEPC
10X Morpholino Propane Sulfonic Acid (MOPS)	41.8 g MOPS 20ml 0.5M EDTA 16.8ml 3M NaOAc DEPC treated water upto 1L. pH 7.00
RNA loading buffer	0.72 ml formamide 0.16 ml 10X MOPS 0.26 ml formaldehyde 0.18 ml DEPC treated water (DPH) 0.1 ml 80% glycerol 0.08 ml BPB 50 μg EtBr
Denaturation mix	13 μl 37% formaldehyde 22 μl formamide 65 μl 10X MOPS buffer

3.9.3. SDS-PAGE

Separating Buffer	18.165 g Tris base 0.4 g SDS distilled water up to 80 ml adjust the pH to 8.8 by HCl distilled water up to 100 ml
Stacking Buffer	6.06 g Tris base 0.4 g SDS distilled water up to 80 ml adjust the pH to 8.8 by HCl distilled water up to 100 ml
10% APS	100 mg ammonium persulfate 1 ml distilled water store at -20 C
Separation Gel	1.5 ml Separating Buffer 1.5 ml 30% acrylamide/bisacrylamide (29:1) 2.895 ml distilled water
before casting the gel, add:	100 ul 10% APS 5 ul TEMED
Stacking Gel	0,875 ml Stacking Buffer 0.56 ml 30% acrylamide/bisacrylamide (29:1) 2.01 ml distilled water
before casting the gel, add:	50 ul 10% APS + 5 ul TEMED

10X SDS PAGE Running Buffer	30.3 g Tris
	144 g Glycine
	10 g SDS + distilled water up to 1 L
4X SDS-Sample Loading Buffer	1 ml Tris-HCl
	0.4 g SDS
	2.3 ml Glycerol
	1 ml EDTA
	4 mg BPB
	0.2 ml Beta-mercaptoethanol
	0.1 ml distilled water
10X Transfer Buffer	30.29 g Tris
	144.1 g Glycine
	1 L distilled water
1X Transfer Buffer before use	100 ml 10X Transfer Buffer
	200 ml Methanol
	700 ml distilled water
10X TBST	24.23 g Tris
	87.66 g NaCl
	10 ml Tween-20
	distilled water up to 800 ml
	adjust the pH to 7.5 by HCl
	distilled water up to 1 L

3.9.4. Culture Media

Dulbecco's modified Eagle's medium (DMEM) and fetal bovine serum (FBS) were purchased from Gibco (Paisley, UK). Penicillin/Streptomycin was commercially obtained from BIOCHROM AG (Berlin, Germany).

10X PBS	80 g NaCl
	2 g KCl
	14.4 g Na2HPO4 • 2H2O
	2.4 g KH2PO4
	distilled water up to 800 ml
	adjust the pH to 7.4 by HCl
	distilled water up to 1 L

3.9.5. Microbiological Media and Antibiotics

Luria-Bertani medium (LB)	10 g tryptophan
	5 g yeast extract
	10 g NaCl
	Distilled water up to 1 L, autoclaved
Ampicillin stock	100 mg Ampicillin
	1 ml distilled water
Kanamycin stock	50 mg/ml in distilled water
	Sterilized by filtration and stored at
	$-20°C$
	50 μg/ml (working concentration)

3.9.6. Staining Solutions

14.3 mM DAPI solution	5 mg DAPI
	1 ml distilled water
before use	dilute with PBS until 300 nM
1 mg/ml PI stain	1 mg PI
	1 ml distilled water
before use	979 μl PBS
	20 μl 1 mg/ml PI stain
	1 μl 10 mg/ml RNase A

3.10. Equipment

Atomic Force Microscope	
Autoclave	Midas 55, Prior Clave, UK
Balances	DTBH 210, Sartorius, GERMANY
	Electronic Balance VA 124, Gec Avery, UK
Cantilevers	PNP TR-20 of NanoWorld (Switzerland) and OTR4 from Bruker (MA, USA)
Carbon dioxide tank	2091, Habaş, TURKEY
Cell culture incubator	Hepa Class 100, Thermo, USA
Centrifuges	Ultracentrifuge J2MC, Beckman Coulter, USA
	Mini Centrifuge 17307-05, Cole Parmer, USA
	Centrifuge 5415R, Eppendorf, USA
	Centrifuge, Allegra X-22, Beckman Coulter, USA
Deepfreezers	−86°C ULT Freezer, ThermoForma, USA
Documentation System	SynGene, UK
Flow Cytometer	BD Accuri C6, USA
Heat blocks	DRI-Block DB-2A, Techne, UK
Hemocytometer	Improved Neubauer, Weber Scientific International Ltd, UK
Laminal flow cabinet	Labcaire BH18, UK
Magnetic Stirrers	M221 Elektro-mag, TURKEY
Micropipettes	Finnpipette, Thermo, USA
Microscopes	Inverted Microscope, Axio Observer Z1, Zeiss, USA
	Confocal Microscope SP5-AOBS, Leica Microsystems, USA
	Stereomicroscope SZ51, Olympus, Japan
Microwave oven	M1733N, Samsung, MALAYSIA

pH meter	WTW, GERMANY
Pipettor	Pipetus-akku,Hirschmann Laborgeräte, GERMANY
Power Supply	Biorad, USA
Real Time PCR	
Refrigerators	2082C, Arçelik, TURKEY
	4030T, Arçelik, TURKEY
Shakers	VIB Orbital Shaker, InterMed, DENMARK
	Lab-Line Universal Oscillating Shaker, USA
Software	FlowJo
	Light Cycler480 Software, Roche, CA, USA
	BD Accuri C6 Software, BD Biosciences, NJ, USA
	ImageJ, NIH
	SynGene G:Box Chemi-XRQ GENESys, India
	GraphPad Software, California, USA
	FlowJo, Oregon, USA
Spectrophotometer	NanoDrop 1000, USA
Thermocyclers	Applied Biosystems 2720 Thermal Cycler
	Applied Biosystems GeneAmp PCR System 2700
	BIORAD DNAEngine Peltier Thermal Cycler
Vacuum pump	KNF Neuberger, USA
Vortex	Vortexmixer VM20, Chiltern Scientific, UK
Water baths	TE-10A, Techne, UK

4. METHODS

4.1. Site Directed Mutagenesis and Deletion of N-Terminus

Alanine substitution was performed on the residues from 292 to 296. The mutagenesis was performed sequentially, one substitution at once. The mutations performed were called as AFTPP, AFTPA, AATPA, AFAPA, AFAAA and AAAAA. For instance, AFTPA code means that 292I and 296P were only replaced with Alanine amino acids. N-terminal deletion mutation was performed from the beginning to the 52nd residue of the open reading frame (ORF).

4.1.1. Polymerase Chain Reaction

In order to mutate 5 residues mutant primers were designed via PrimerX tool of Bioinformatics.org (Table 3.1) annealing to the site-of-mutation. The melting temperatures were predicted in the online tool of NEB TM Calculator (version 1.9.10). PCR amplification of mutant 5' or 3' CDS was performed by Phusion High-Fidelity DNA Polymerase (New England Biolabs, MA, USA) and the reactions were prepared according to Table 4.1 and run according to the conditions given in Table 4.2.

Table 4.1: PCR reaction reagents and amounts.

Reagent	**Final Concentration**	**Amount**
CamAlstR-C ORF template		1 μl
5X Phusion HF Buffer	1X	4 μl
10 μM Forward Primer	0.2 μM	0,4 μl
10 μM Reverse Primer	0.2 μM	0,4 μl
dNTPs (2.5 mM each)	50 μM each	0,4 μl
Phusion DNA Polymerase (2U/ μl)	0,4 U	0,2 μl
100% DMSO	20%	0,4 μl
Distilled water		13,2 μl

Table 4.2: PCR reaction conditions.

Initial Denaturation	98 °C	30 sec	X 1
Denaturation	98 °C	10 sec	X 30
Annealing	(calculated melting temperature)	15 sec	
Extension	72 °C	20 sec	
Final Extension	72 °C	7 min	X 1

The same reactions were repeated for the amplification of the total CDSs with the mutation. In order to obtain full length CDS, products of each 3' and 5' products were combined in the same PCR reaction tube as the templates. CDS was amplified by using the ATG and Stop primers. For N-deletion, primer was designed to be having an artificial ATG codon and annealing to the sequence after 52nd amino acid codon of the receptor.

4.1.2. Agarose Gel Electrophoresis and DNA Extraction

PCR products were run in 1% agarose gel in 90V for 30 min and extracted from the gel by using Agarose Gel DNA Extraction Kit (Sigma-Aldrich, Germany) as recommended in the kit protocol.

4.1.3. Cloning

Amplified and extracted DNAs were cloned into previously modified pcDNA3/HA vector. The purified vector and insert DNA were cut by HindIII (Thermo Fisher Scientific, MA, USA) and EcoRI (Thermo Fisher Scientific, MA, USA) enzymes. The restriction reaction mixture was prepared by mixing 1 μg of DNA, 1 μl of each enzyme, 2 μl of 10X reaction buffer and distilled water up to 20 μl. The mixture was incubated at 37 °C for 30 min. Enzyme inactivation was performed at 65 °C for 15 min. The cut products were run in 1% agarose gel electrophoresis and extracted. Ligation reactions were performed by using T4 DNA Ligase (New England BioLabs, MA, USA). The cut

insert and vector samples were mixed in a 1:3 ratio. 1 μl T4 DNA ligase, 2 μl 10X T4 DNA Ligase Bufer and distilled water were added until the total volume was 20 μl. The mixture was incubated at RT for 20 min. Enzyme inactivation was performed at 65 °C for 10 min. And the product was diluted with 80 μl distilled water. 20 μl of this ligation reaction was mixed with previously prepared competent DH5α bacteria. Transformation was performed via the steps as follows: 1) incubation in ice for 50 min 2) Heat shock at 42 °C for 45 sec 3) incubation in ice for 2 min 4) 1 ml LB medium was added and 5) incubation at 37 °C for 1hr on a shaker ($\geq$200 rpm) 6) spread the transformed bacteria onto Ampicillin containing LB agar plate 7) left for incubation O/N at 37 °C. The colonies were selected and used for colony PCR. Colony PCR reactions were prepared as given in Table 4.3 by using Taq DNA Polymerase (New England BioLabs, MA, USA). The colonies were added directly into the mixtures. The reaction conditions are given in Table 4.4.

Table 4.3: Colony PCR reagents and amounts.

Reagent	**Final Concentration**	**Amount**
10X Standard Taq Reaction Buffer	1X	2 μl
10 μM Forward Primer	0.2 μM	0,4 μl
10 μM Reverse Primer	0.2 μM	0,4 μl
dNTPs (2.5 mM each)	50 μM each	0,4 μl
Taq DNA Polymerase (2U/ μl)	0,4 U	0,2 μl
100% DMSO	20%	0,4 μl
Distilled water		16,2 μl

Table 4.4: Colony PCR conditions.

Initial Denaturation	95 °C	30 sec	X 1
Denaturation	95 °C	15 sec	X 25
Annealing	(calculated melting temperature)	15 sec	
Extension	68 °C	90 sec	
Final Extension	68 °C	7 min	X 1

The colony PCR products were run in 1% agarose gel electrophoresis. The positive colonies were selected and grown in ampicillin containing liquid LB medium at 37 °C O/N. Isolation of plasmids from positive colonies were performed with the NucleoSpin Plasmid Kit (Macherey-Nagel, UK). The purified plasmids were sent to Macrogen Korea for sequencing and presence of the mutations was verified. For generation of stable cell lines, CamAlstR-C CDS was digested and inserted into pLENTI-III-HA vector with KpnI and BamHI restriction enzymes. For control cell lines, mCherry sequence was digested from mCherry2-N1 plasmid and inserted into pLENTI-III-HA vector with ApaI and NotI enzymes.

4.2. Preparing Fluorescently Tagged Receptors

In order to tag CamAlstR-C, previously prepared pcDNA3-SYFP2 vector was used. The receptor was amplified with the WT-wo-stop primer (Table 3.1) to delete the stop codon. Cloning of this part was the same as in Section 4.1.3.

4.3. Studies on Mammalian Cell Culture

The cell lines in analyses were given in Table 4.5. All cell lines were grown in Dulbecco's Modified Eagle Medium (DMEM by Gibco) containing 10% Fetal Bovine Serum (FBS) and 1X Penicillin-Streptomycin (Gibco). The cells were incubated in 5% CO_2 conditioned 37 °C incubator in 10 cm cell-culture Petri plates. All treatments were performed under a sterile hood. The passages were performed every 3-4 day with a 1:5 dilution. The medium was removed and the cells were washed with 1X PBS twice. PBS was removed and the cells were detached by adding 1 ml 2.5% trypsin-EDTA solution. The cells were incubated at 37 °C for 2 min. For inactivation of trypsin complete DMEM was added in a volume at least 3 times of trypsin and the cells were centrifuged at 600 g for 4 min in a falcon tube. The medium was removed and the cells were resuspended in complete growth medium. Complete growth medium was added onto a clean 10 cm plate and the required amount of cell suspension was added dropwise. For the storage of cells, the suspended cells were taken into cryovials and DMSO was added for a final concentration of 10% and the cryovials were put into

isopropanol containing freezing box at -80°C refrigerator. For long term storage the tubes were put into -150 °C.

Table 4.5: The list of cell lines used in different analyses.

Cell Line	**Type**	**The Analysis Performed**
Huh7	Hepatocellular carcinoma	AFM, IF, Xenograft, XTT
293FT	Human embryonic kidney	XTT
HepG2	Hepatocellular carcinoma	XTT
HeLa	Ovarian cancer	XTT
MCF7	Breast cancer	XTT
MeWo	Melanoma cells	XTT

4.3.1. Transient Transfection

Huh7 cell line was supplied by Mehmet Öztürk's laboratory (Bilkent University, Ankara). Transient transfection was performed by using FuGENE HD Transfection Reagent (Promega, WI, USA) in a 3:1 (reagent volume:DNA amount) ratio. Reagent-DNA mixture was incubated at RT for at most 15 min. This mixture was added onto the cells and left for incubation at 37 °C for 24-48 h. Each transfection was verified for efficiency with an EGFP containing control plasmid (pEGFP-N2) and for transfection toxicity with an empty plasmid (modified pcDNA3/HA). For the AFM studies the cells were transfected with pcDNA3/HA-CamAlstRC plasmids. On the other hand for immunofluorescence studies, they were transfected with pcDNA3-SYFP2-CamAlstRC plasmids.

4.3.2. Generation of Stable Cell Lines

Antibiotic kill curve experiments were performed on Huh7 cells with various concentrations of Puromycin (0, 250, 500, 1000, 1500, 2000, 5000, 10000 ng/mL). The cells were seed on 24-well plates in duplicates. Antibiotic selection was performed for one week. The minimum concentration of Puromycin that killed all cells in the well

was determined as 2 μg/mL. So, for the stable cell selection this concentration was chosen. And for the maintenance of cells after selection process, 1 μg/mL was chosen. For virus collection, 293FT cells were transfected with pLENTI-III-HA vector system having two packaging vectors (ABM Inc., Canada). CamAlstR-C was cloned into pLENTI-III-HA vector with the protocol described in Section 4.1.3. Transfection was performed with the help of cationic reagent Turbofect (Thermo Fisher Scientific, MA, USA). The growth medium was replaced 4 hr after transfection. The same transfection was performed with pLENTI-III-mCherry and empty pLENTI-III vectors, again together with the packaging vectors. Two days after transfection the growth media on the cells were collected into clean falcon tubes. The tubes were centrifuged at 400 g for 4 min to get rid of dead cells. The supernatant was filtered through 0.45 μm filter. At last the filtered media were taken into aliquots and stored at -80 °C. Transduction was performed on Huh7 cells. 500.000 cells were seed on 6-well plates one day before transduction. Virus containing medium was mixed with total growth medium in 1:1 ratio and 0.1% polybrene was added. This mixture was added onto cells. One well of cells was not transduced as antibiotic-kill control group. The media were replaced 6-8 hr after transduction. Four days after transduction, the cells were treated with 2 μg/mL Puromycin. This selection was continued by replacing the medium everyday with the same amount of Puromycin, for one month.

4.3.3. RT-PCR Verification of Expression in Stable Cells

Stable cells were lysed in RLT buffer of QIAGEN RNAeasy Kit. Following RNA isolation procedure was performed according to manufacturer's recommendations. RNA concentrations were measured in nanodrop. cDNA synthesis was performed with oligo-dT primer of ImProm II Reverse Transcription System (PRomega, WI, USA) according to manufacturer's recommendations. This cDNA mix was used as the template in PCR reaction performed with CamAlstRC forward and reverse primers given in Table 3.2. The expected product of these primers would be 1357 bp.

4.3.4. SDS-Polyacrylamide Gel Electrophoresis (PAGE) and Western Blotting

RIPA buffer (with protease inhibitors) was directly added onto wells and the plate was incubated on ice for 30 min. Cells were scraped and taken into Eppendorf tubes. The suspensions were homogenized with syringes and then centrifuged for 10 min. at maximum speed. The samples were mixed with loading dye and denatured at 90 °C for 5 min.SDS-PAGE gels were prepared, run and transferred using Mini-Protean III cell and Mini Trans-blot cell (BioRad). 10 % running gel was prepared. 5% stacking gel was prepared onto it. Protein samples were prepared in 1X SDS-PAGE dye. 5 μl of pre-stained protein marker (Thermo Scientific Inc.) was used as the molecular weight standard. The gels were run at 80-100V at first and then the voltage was increased up to 120V. Transfer was performed onto PVDF membrane (Thermo Scientific Inc.). For the transfer, the tank was run for 2 hr at 100V in cold-room. After transfer, blocking was performed in 5 % non-fat milk for 1 hr at room temperature. The membrane was not washed. Primary antibody incubation was performed at 4°C overnight. The membrane was washed with TBST for three times. Secondary antibody incubation was performed 1hr at room temperature. Then, membranes were washed twice TBST. For HRP imaging, West Pico Chemiluminescent Substrate was used. Finally, ImageJ analysis software was used to quantify band intensities. Primary antibodies were used as 1:5.000 anti-HA (Sigma), 1:2000 anti-GAPDH (Santa Cruz). For the secondary antibodies 1:5000 HRP conjugated anti-rabbit or anti-mouse antibodies were used.

4.3.5. Immunofluorescence

pcDNA3-SYFP2-CamAlstRC transfected cells were washed with PBS twice. Ice-cold 4% paraformaldehyde was added onto the cells and the plate was incubated for 20 min on a shaker in RT. Washing was performed with PBS three times for 5 min. Blocking was performed in 2% BSA solution for 30 min on a shaker in RT. The cells were incubated with anti-ZO1 antibody (Invitrogen, catalog number: 40-2200) in a 1:200 ratio at +4 °C for overnight on a shaker. Washing was performed with PBS three times for 5 min. The cells were incubated in secondary antibody (Alexa 555-

conjugated α-rabbit IgG, 1:200 ratio) for 4 hr at RT. Washing was performed in PBS three times for 10 min. The cells were added 300 nM DAPI solution and incubated at RT for 2 min. Then washing was performed for twice with PBS. The coverslips were taken onto clean slides and covered with PBS. These samples were observed under confocal microscope.

4.3.6. Atomic Force Microscopy

Atomic force microscopy experiments were performed in BioAFM Laboratory of Hamdi Torun in Electrical and Electronical Engineering Department of Boğaziçi University. PNP TR-20 (k = 80 pN/nm, NanoWorld) and OTR4 silicon nitride cantilever (k = 80 pN/nm, Bruker) were used. The cantilevers were incubated with 100 ng/ml AST-C peptide for 20 min at RT and washed with 1X PBS. For AST-A binding experiments OTR4 silicon nitride cantilever (k = 80 pN/nm, Bruker) was incubated in 1 mg/ml Allatostatin IV peptide for 15 min at RT. A few drops of 1X PBS was added onto the cells and the sample coverslips were taken onto sample stage. Laser was positioned as it strikes the cantilever end, before the sample was placed. After sample placement, piezoelectric translator head was lowered in scanning tunneling microscope (STM) mode until a water meniscus was formed between peptide solution and PBS on the cells. Than the distance between cantilever tip and sample was adjusted to 1 mm via focusing of a single cell. The Sum signal was maximized here and the piezoelectric translator head was allowed to engage with a speed of 0.25 μ/step. Set point was adjusted for 0.3 V. On each sample more than 50 recordings were counted. For each sample data were collected on different loading rates (0.1 μ/s, 0.5 μ/s, 1 μ/s, 3 μ/s, 5 μ/s, 10 μ/s, 15 μ/s, 20 μ/s, 30 μ/s and 40 μ/s). Each WT and mutant receptor recordings (AFTPP, AFTPA, AFAPA, AATPA, AFAAA, AAAAA and Ndel) were performed in biological duplicates. The experiment was calibrated in an empty area with no cell. The control groups were set as a sample with no CamAlstR-C expression (empty pcDNA3 transfected cells), a sample which cantilever penetrated the cell membrane and another sample washed with excess peptide to saturate CamAlstR-C receptors. The unbinding events were evaluated against non-adhesive events and these

were taken for force calculations. The AFM apparatus measured the laser deflection coming from cantilever tip as voltage units and we produced force-distance (FD) curves from potential-distance curves, according to Eq. 4.1

$$F = k\delta z \tag{4.1}$$

Here, k is the spring constant of the cantilever and δz is the cantilever deflection change in z axis. This deflection was calculated by Eq. 4.2 from calibration curve:

$$\delta z = Vtot\ 1/S \tag{4.2}$$

Here, S is optical lever sensitivity and obtained from the slope of linear part of calibration curve on a stiff surface which is assumed to be "indefinitely hard". Vtot is the potential difference when the cantilever moves one step towards and against the sample. Finally, the force differences in unbinding events were assumed as the interaction forces between our receptor and its' ligand. After converting the raw data into FD curves, histograms were generated for detection of most probable rupture forces. These rupture forces were plotted against loading rates in dynamic force spectrum and

the dissociation rate constants were calculated via Eq. 4.3

$$\text{Koff} = \text{ro} \,/\, \text{f}\beta \tag{4.3}$$

Here r0 is the loading rate (pN/s) at zero force and fβ is the slope of the curve of dynamic force spectrum.

4.3.7. Cell Cycle Analysis

Cells were trypsinized, centrifuged and washed with PBS twice. Then centrifugation was performed at 500 g for 5 min. The pellet was resuspended in 1 mL PBS. 70% ice cold EtOH was added dropwise. The cells were centrifuged was resuspended in 2 mL PBS. Propidium iodide (PI) staining was performed with a concentration of 10 μg/mL for 1 min in dark. The cells were washed with PBS, centrifuged and resuspended in PBS. The samples were analyzed in Accuri BD FACS system.

4.3.8. Viability Assay

For XTT cell viability assay Cell Proliferation Kit II (Sigma-Aldrich, Germany) was used. The cells were seed into a 96-well plate in various numbers depending on the cell line, in triplicates. Next day various concentrations (100 nM, 1 μM, 10 μM, 50 μM AST-C and 0.1%BSA for control) of AST-C peptide were added onto wells. Treatments were performed for both the peptide that was used in previous experiments (*Drosophila* AST-C, DroAST-C) and the peptide that was not used before and predicted from RNAseq data (*C. morosus* AST-C, CamAST-C). Different untransfected cancer cell

lines originating from different tissues were used (Table 4.5). The cells were incubated at 37°C and 5% CO_2 for 48 hr. The complete growth medium (including phenol red) was replaced with complete DMEM F12 medium (without phenol red). XTT reagents were mixed in 1:50 ratio and added onto wells. The reagents were incubated with the cells for 2 hr. Measurements were taken at both 490 nm and 650 nm. Each experiment was performed in at least three biological replicates.

4.4. *in vivo* Mouse Studies

4.4.1. Animals

The project protocol was approved by the Animal Ethics Committee of Boğaziçi University Vivarium. 2-3 months old male NUDE mice were used. All mice were kept in IVC cages in 23 ± 2 °C rooms, with a 12-hr-day-night cycle. The animals were fed *ad libitum*.

4.4.2. Xenograft

Stable cells were grown in 15 cm plates in complete growth medium including maintenance Geneticin concentration (500 μg/mL). The cells were trypsinized and washed with PBS twice. Finally, approximately 4 million cells were suspended in 200 μl PBS and grafted into both flank regions of male NUDE mice subcutaneously. Xenograft injections were performed under isofluoran anesthetics conditions. For the first experiments, 3 mice were used. AlstR-stable cells were grafted into the left flank region while mCherry-stable cells were grafted into the right flank region. For the second experimental set-up, all AlstR-stable, mCherry-stable and untransfected Huh7 cells were grafted on different combinations to the animals.

4.4.3. AST-C Treatment on Tumors

AST-C peptide of *Manduca sexta* was dissolved in 0.1%BSA in PBS for stocks. 10 μM AST-C dilutions were prepared only in PBS in order to use in mouse treat-

ments. For control mice, Saline injections were done. The treatments were performed subcutaneously in close proximity to tumor region. For some animals intra-tumoral injections were performed. The tumors on mice were measured with a caliper every day. When the size of the tumors have reached the ethical limit (sum of two tumors $\leq$ 3cm), the mice were sacrificed, tumors were weighed and stored at -80 °C.

4.4.4. Live Imaging

Live images of mCherry-stable tumors were taken in IVIS Spectrum in vivo Imaging System. The mice were subjected to isofluorane anesthesia and fluorescent images were acquired via transillumination. Radiant efficiency and the area of the region of interest (ROI) were measured before treatment and some days after treatment.

4.5. Sampling of Stick Insects

The stick insects (*C. morosus*) were obtained from University of Cologne, Germany. The animals were held in cages at room temperature and fed *ad libitum*. The adult females were sacrificed via CO_2 and cooled down in PBS in -20 °C. Dissection was performed in presence of cold PBS. The organs were immediately put into Trizol reagent and stored at -80 °C. For the total body RNA isolation, adult animals were directly sacrificed in liquid nitrogen and disrupted by a mortar and pestle in presence of liquid nitrogen. Then the body samples were stored at -80 °C.

4.5.1. RNA Isolation

Total body samples of stick insects were weighed and further disrupted by a mortar and pestle in liquid nitrogen. 1 mL Trizol was added and the mixture was homogenized in MagnaLyzer. The dissected organs were filled with Trizol up to 1 mL and homogenized by a manual motor pestle. The homogenates were taken into clean tubes and incubated at room temperature for 5 min. Chloroform was added in a 1:5 Trizol:Chloroform ratio, vortexed for 20 sec and incubated at room temperature for 2 min. Centrifugation was performed at 10000 g for 18 min. The aqueous part was taken

into a clean microcentrifuge tube and 1 volume of 100% EtOH was added. The tube was inverted for 6 times. 700 μl of this sample was loaded into NucleoSpin® RNA Column. Centrifugation was performed at 11000 g for 30 sec and the flowthrough was discarded. The following procedure was performed as recommended in the protocol of NucleoSpin® RNA (740955.50, MN, Germany). DNase treatment was performed as recommended in the same protocol. For shipping to USA, the samples were dried in RNAstable (Biomatrica).

4.5.2. MOPS Gel Electrophoresis

All RNA analyses were performed under RNase –inhibiting conditions with the help of DEPC-treated water and RNase-ZAP (Sigma-Aldrich, Germany). MOPS gel was prepared in 0.8% ratio. 400 ng of RNA sample was denatured in denaturing mixture at 55 °C for 15 min. The samples were mixed with 4X loading dye and run in gel at 60 V for 40 min. The intact RNA samples were chosen for sequencing and expression analyses.

4.5.3. Semiquantitative RT-PCR

The primers were designed from the putative GPCR transcripts via BlastPrimer tool [30]. The list of primers was given in Table 3.2. In order to detect the presence of gDNA contamination, MyTaq One-Step RT-PCR kit (Bioline, London, UK) was used and one reaction from each RNA sample was prepared without addition of reverse transcriptase enzyme (no-RT reaction). The RNA samples which did not yield any product in no-RT reaction were included in the following reactions as pure RNA samples. Amount of RNA of different tissues was adjusted to 1 μg. First strand cDNA synthesis was performed as recommended in the protocol of SensiFAST cDNA Synthesis Kit (Bioline, London, UK). The reaction ingredients were mixed as given in Table 4.6. The mixture was prepared in ice. The initial incubation was performed at 25 °C for 10 min, then 42 °C for 15 min and finally 85 °C for 5 min. The cDNA samples were freshly used in PCR reactions and stored at -20 °C for later use.

Table 4.6: The ingredients of cDNA synthesis reaction.

Reagent	**Final Concentration**	**Amount**
RNA		1 μg
5X TransAmp Buffer	1X	4 μl
Reverse Transcriptase		1 μl
Nuclease free water		up to 20 μl

For PCR, every primer couple was optimized with regard to annealing temperature. The most specific primers were chosen for analysis. PCR mixes were prepared according to the amounts given in Table 4.7and the conditions were the same as given in Table 4.4

Table 4.7: Reagents and amounts for RT-PCR.

Reagent	**Final Concentration**	**Amount**
cDNA sample		0.5 μl
5X Phusion GC Buffer	1X	4 μl
10 μM Forward Primer	0.2 μM	0,4 μl
10 μM Reverse Primer	0.2 μM	0,4 μl
dNTPs (2.5 mM each)	50 μM each	0,4 μl
Phusion DNA Polymerase (2U/ μl)	0,4 U	0,2 μl
100% DMSO	20%	0,4 μl
Distilled water		13,7 μl

RT-PCR products were run in 1% agarose gel at 90 V for 30 min and the images were quantified in ImageJ with regard to GAPDH values.

4.5.4. qPCR

Quantitative PCR reactions were prepared according to the recommendations of SensiFAST SYBR No-ROX Kit (Bioline, London, UK). Each sample was prepared in

technical duplicates. The reaction mixtures were prepared according to the Table 4.8 in a 96-well plate. Reaction conditions were given in Table 4.9. The reactions were run in Light Cycler 480 (Roche Molecular Systems, CA, USA). ΔΔCp values were calculated and plotted, with regard to GAPDH expression and using the least expressed organ as the calibrator. Each qPCR analysis was performed in biological triplicates, each having at least three animals.

Table 4.8: The reagents of qPCR mixture.

Reagent	**Final Concentration**	**Amount**
2X SensiFAST SYBR No-ROX Kit	1X	10 μl
10 μM Forward Primer	0.4 μM	0,8 μl
10 μM Reverse Primer	0.4 μM	0,8 μl
cDNA		0,5 μl
Nuclease free water		7,9 μl

Table 4.9: The reaction conditions for qPCR.

Polymerase Activation	95 °C	2 min	X 1
Denaturation	95 °C	5 sec	X 40
Annealing	60 °C	10 sec	
Extension	72 °C	20 sec	
Melting Curve	from 55 to 95 °C		X 1
Cooling	24 °C		

4.6. Bioinformatics Studies

4.6.1. Molecular Dynamics Simulations

Molecular dynamics studies were performed via NAMD on Visual Molecular Dynamics (VMD) software utilizing both AST-A and AST-C peptides. Coordinate files were obtained from the model generation results of previous thesis [31]. The simula-

tions were run for 10 ns with 2000 minimization steps at 298 K. Each simulation was performed in triplicates.

4.6.2. Sequencing and *de novo* RNA Assembly

2 x 100 bp paired-end sequencing was generated by GENEWIZ Inc. (NJ, USA) using their own protocol. Assembly was performed by GENEWIZ Inc. using Trinity. Sequence files were retrieved in fastq format. *Blatella germanica* and *Zootermopsis nevadensis* were chosen as the closest genomes to *C. morosus*. The data was mapped onto *B. germanica* and *Z. nevadensis* genome using TopHat 2.1.1 tool. Functional and structural annotations were performed via Blastx against NCBI *B. germanica* sequences. The data was be submitted to Sequence Read Archives (SRA).

4.6.3. Neuropeptidome Analysis

All known neuropeptide precursor sequences were taken as query and used in tblastn tool. In the search set parameters of tblastn, assembled Trinity dataset was chosen. Each resulting transcript was checked in online blastx tool. Corresponding ORF of the transcript was taken and searched for signal peptide sequence in SignalP 4.1 [32]. Proprotein convertase cleavage sites were predicted according to the rules mentioned in Duckert *et al.* [33]. Disulfide bonds were analyzed in DIANNA server [34]. Other posttranslational modifications such as sulfation states of tyrosine residues, cyclization of N-terminal glutamine/glutamic acid residues or C-terminal amidation states were also checked *in silico*. Isoforms of the corresponding transcripts were analyzed with the same steps. Additionally, the resulting mature peptides of these transcripts were aligned for comparison.

4.6.4. GPCRome Prediction

The ORFs were analyzed for the presence of transmembrane helices (TM) in TMHMM [35]. TM containing ORFs were separated into different files. Most of the one TM-containing ORF were due to signal sequences, so that these were excluded.

The ORFs which have more than 2 TM regions were taken into alignment with NCBI Blastp tool [36]. The sequences were aligned on NCBI server within non-redundant protein database without any other restriction. Results of the top ten hits were taken and filtered according to presence of GPCR domains. Finally, putative GPCR ORFs were separated into 5 classes according to the nomenclature used by GPCRdb [37].

4.7. Statistical Analyses

In AFM results, the most probable rupture forces were obtained from the modes of the histograms and the errors were calculated as Standard Error of the Mean. In XTT viability assay graphs, each experiment was performed at least in three biological replicates and statistical errors were calculated as Standard Error of the Mean. In qPCR and semiquantitative RT-PCR results the graphs were plotted with the standard deviation of biological replicates and the multiple comparisons were calculated via Two-Way ANOVA method with Tukey test to see the intra-group variations. To compare the patterns of qPCR and semiquantitative RT-PCR results, Spearman correlation test was performed.

5. RESULTS

5.1. IXTPP Motif on ECL3 is Important for Ligand Binding

The first GPCR from *C. morosus* has been identified in our laboratory and named CamAlstR-C, previously [31]. In the same study, a putative ligand binding motif (IXTPP residues) was proposed on the ECL3 (Figure 1.1). Then we continued to support the importance of these residues via ligand binding studies. Mutagenesis was performed on the wild type (WT) receptor and five amino acids (from 292I to 296P) were substituted with Alanine residues. *In silico* analyses showed that the system could be stabilized between the WT receptor and AST-C peptide in 10 ns of MD simulations (Figure 5.1a). However, in the same analyses, it was slightly unstable between the WT receptor and AST-A (Figure 5.1b), between AST-C peptide and the Ndel receptor (Figure 5.1c). Even the distance between the α carbons of the interacting residues increased in any of the mutant receptors (Figure 5.1d).

In order to evaluate the importance of IXTPP motif, the receptor-ligand binding was analyzed on living cells in AFM experiments. Huh7 cells were transfected with CamAlstR-C and it was localized to the membrane in IF controls (Figure 5.2).

These cells were utilized in AFM experiments. Interpretation of the graphs is illustrated in Figure 5.3 and the experimental set-up in Figure 5.4. The cell is attached on a solid surface and the peptide is bound to the tip of the cantilever. As shown in Figure 5.3, until the cantilever approaches to the cell surface (Stage I) the force on the cantilever increases slightly with the drag force of the liquid environment. On the surface of the cell (Stage II), the repulsive force increases sharply during the approaching movement and decreases during the retraction movement. Just after the cell surface level during retraction movement (Stage III) unbinding events occur. And these specific events are evaluated as rupture force values. In our experimental set-up, empty pcDNA3/HA transfection was performed for control experiments (Figure 5.4a) because allatostatin receptors are not expressed in mammalian cells. Within 100 approaching-

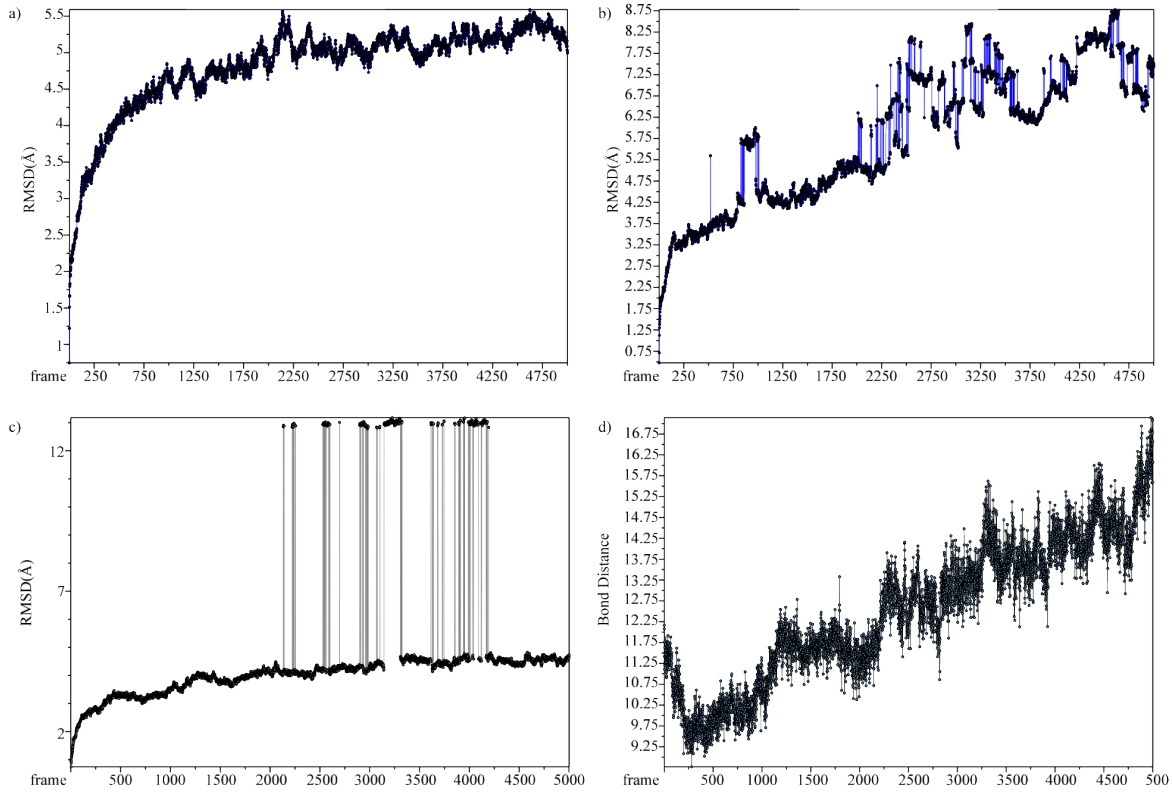


Figure 5.1: RMSD graphs of the total system during 10 ns MD simulations on a) WT CamAlstR-C with AST-C, b) WT CamAlstR-C with AST-A, c) Ndel AlstR-C with AST-C, d) and the change in bond distance between the Cα of A292 on AFTPP mutant and F15 on AST-C peptide, during the same simulations.

retraction movements there were very few (nearly 0) specific binding events. However, when we express WT CamAlstR-C on the same cells, we detected an average of 25 unbinding events within 100 retraction movements (Figure 5.4b). For the saturation control, we washed the cells with excessive AST-C peptide and the unbinding events disappeared again (Figure 5.4c). Therefore, we conclude that AST-C could specifically interact with WT CamAlstR-C in AFM experiments.

Comparing the WT receptor with the mutant forms, AFM results showed that mutant receptors were interacting with lower unbinding forces, barrier widths (xβ) and higher Koff values than the WT receptor (Figure 5.5 and Table 5.1). In addition, N-terminal deletion leads to a decrease in the unbinding forces and increase in the Koff

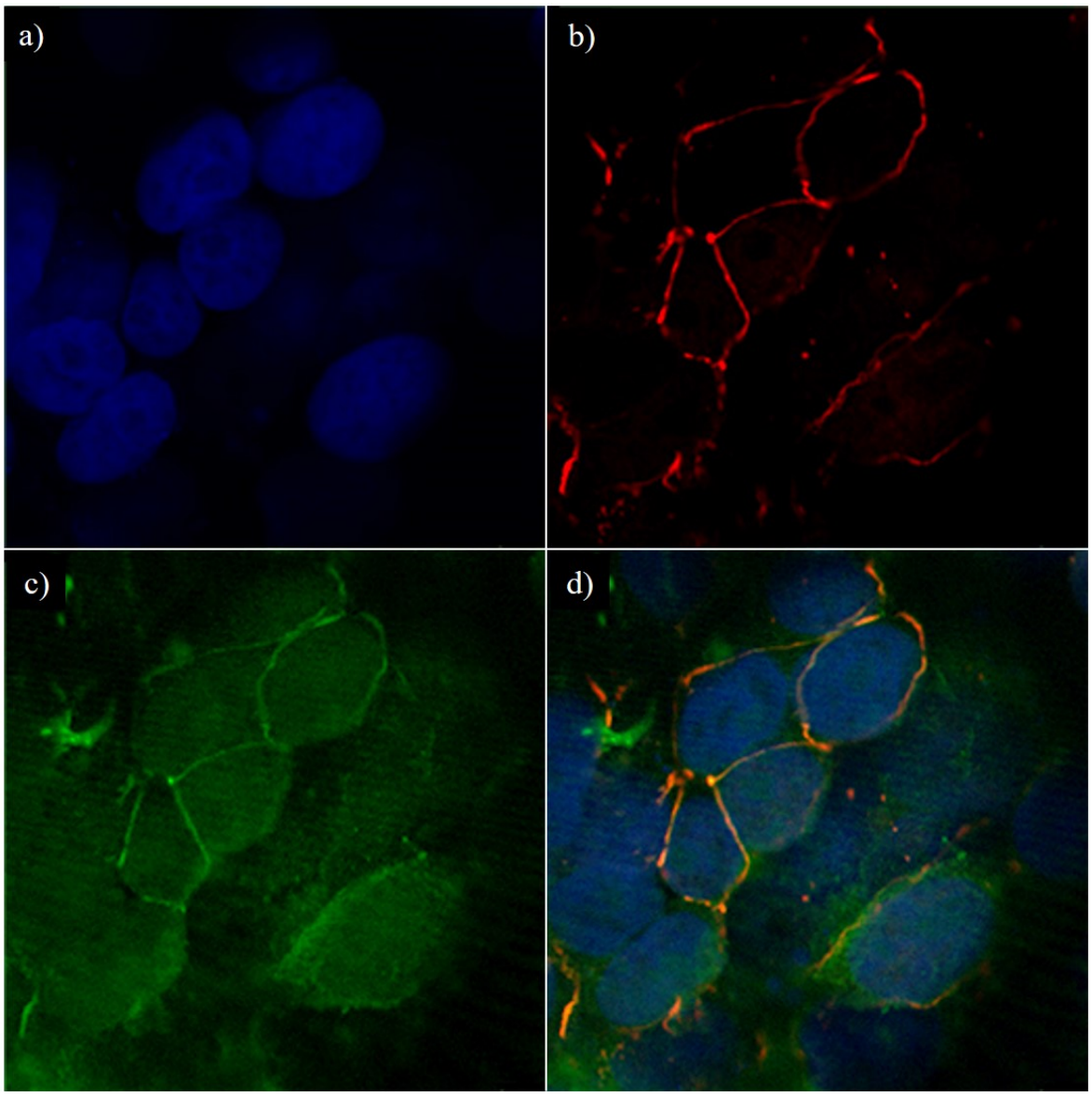


Figure 5.2: CamAlstR-C was localized to the membrane on Huh7 cells. The images were taken under Leica Confocal Microscopy by a) DAPI staining, b) anti-ZO1 staining, c) SYFP2-CamAlstRC and d) from the overlay of channels in a, b and c. The figure was adapted from [7].

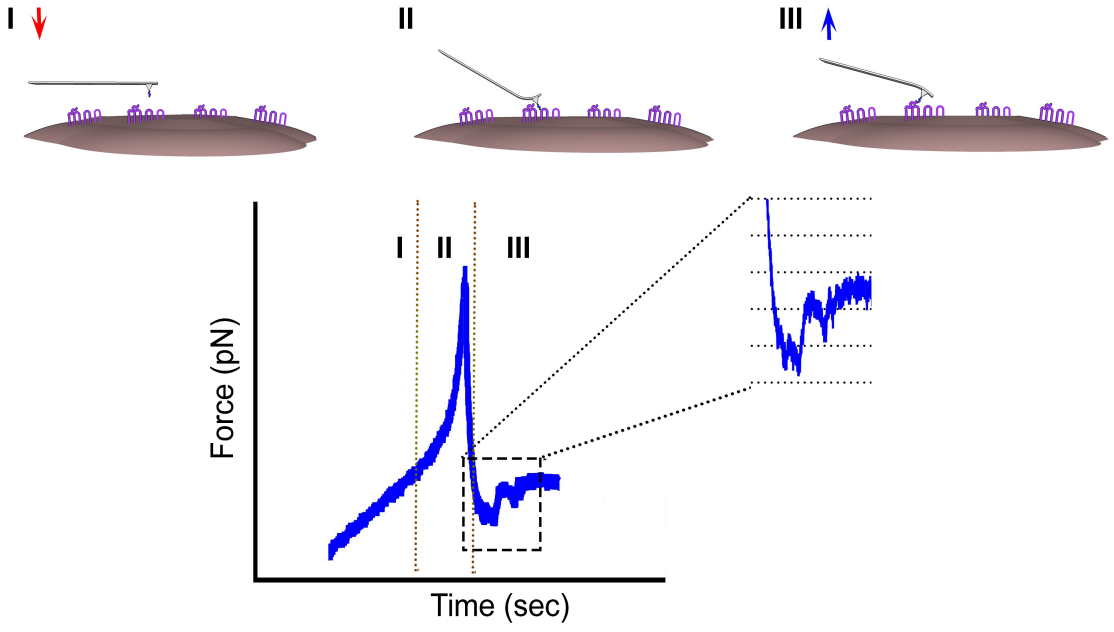


Figure 5.3: The illustration for interpretation of AFM graphs for specific unbinding events. Example for the force-time curve was given with the cell and cantilever illustrations as well as the tilting of the tip of cantilever, at different approaching and retraction steps. The figure was adapted from [7].

values (Figure 5.5 and Table 5.1). Bell's parameters for each mutant receptor showed that the strength of interaction decreased in all of the mutant receptor forms. Koff and xβ parameters for Ndel, AAAAA, AFAAA, AATPA, AFAPA, AFTPA and AFTPP were given in Table 5.1.

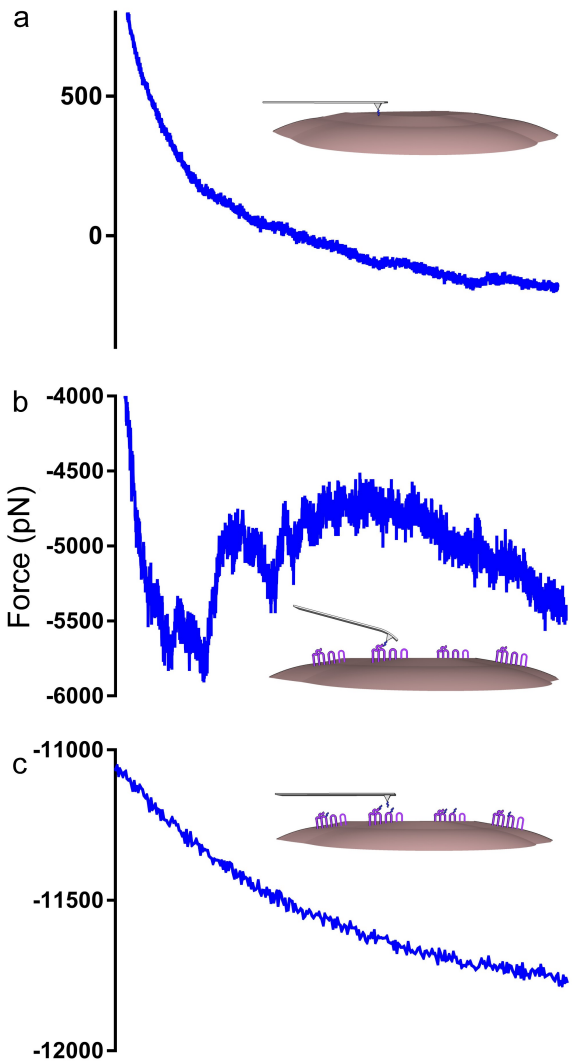


Figure 5.4: The illustration for the experimental set-up in AFM experiments. The force-time graphs were given with the illustrations of a) the empty pcDNA3/HA transfected cells, b) WT CamAlstR-C transfected cells and c) WT CamAlstR-C transfected cells after washing with AST-C peptide. The figure was adapted from [7].

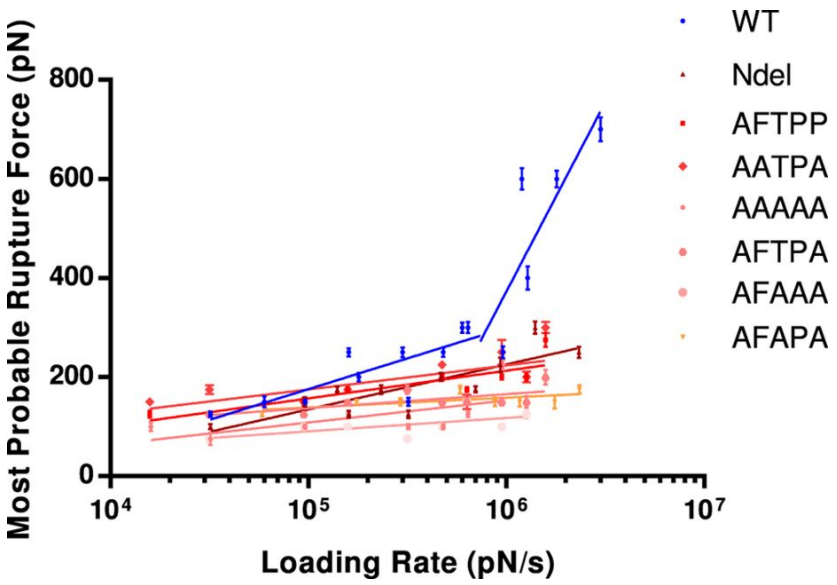


Figure 5.5: Dynamic strength of interactions between AST-C with mutant forms was compared with that of WT receptor in AFM experiments. Blue: WT receptor. The other data points in the legend were colored. Error bars (s.e.m.) were calculated as the mean error of force histograms (n>20). The figure was adapted from [7].

Table 5.1: Bell's parameters for interaction of AST-C with different forms of CamAlstR-C.

Forms of CamAlstR-C	$\mathbf{K}_{off}(\mathbf{s}^{-1})$	$\mathbf{x}\beta$ **(Å)**
WT (at low loading rates)	2.00E+10	0.828
WT (at high loading rates)	3.33E+0.9	0.138
Ndel	2.50E+10	1.040
AAAAA	5.00E+11	2.070
AFAAA	1.00E+11	4.140
AATPA	5.00E+10	2.070
AFAPA	1.11E+11	4.600
AFTPA	1.00E+11	4.140
AFTPP	5.00E+10	2.070

5.2. Both AST-A and AST-C Interact With CamAlstR-C, But With Different Affinities

The difference between the binding forces of AST-A and AST-C with the receptor was shown before [31]. In this study, we have replicated the experiments for various loading rates. Both AST-C and AST-A peptides interacted with CamAlstR-C but with different affinities. AST-C resulted in an increase in the rupture forces by increasing loading rates, but the equation changed in high loading rates (Figure 5.6a), which meant a second energy barrier in this unbinding event (Figure 5.6b). In smaller loading rates its Koff and xβ were calculated as 2.00E+10 s^{-1} and 0.828 Å, respectively. In higher loading rates, these values were 3.33E+09 s^{-1} and 0.138 Å, respectively (Table 5.1). However, this two-step process disappeared in mutant receptor interaction (Figure 5.5).

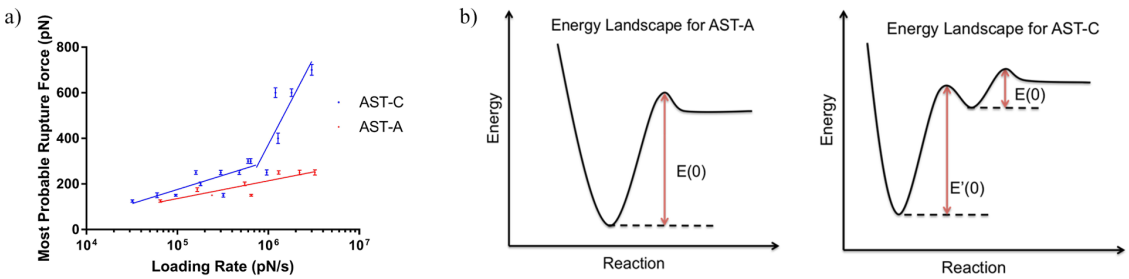


Figure 5.6: The interaction of AST-C differs from AST-A. a) The most probable rupture forces between the receptor and the peptides (N¿50 unbinding events). Errors are calculated via the s.e.m. method. The figure was adapted from [7]. b) Illustration of energy landscape of unbinding events between the peptides and the receptor.

5.3. Neuropeptidome Analysis Revealed CamAST-C Peptide

During the previous AFM studies, the AST-C peptide of *Drosophila melanogaster* was utilized due to the absence of information about *C. morosus* peptide. Therefore we planned to perform RNA sequencing from the total body of *C. morosus*. As a result of RNA sequencing, about 94.820.114 base reads were obtained which contributed to 128.397 assembled transcripts.

We searched the databases for previously identified insect neuropeptides. In total, 65 putative neuropeptide precursor sequences were taken as queries from *D. melanogaster*, *Tribolium castaneum*, *Manduca sexta*, *Homarus americanus*, *Daphnia pulex*, *Tigriopus californicus*, *Schistocerca gregaria*, *Procambarus clarkia*, *Locusta migratoria*, *Ceratitis capitate*, *Nilaparvata lugens* and *Culex quinquefasciatus* (Appendix C). We could identify the transcripts for 23 of these neuropeptide precursors (Table 5.2). They contributed to 29 of the assembled transcripts. We could detect signal peptide sequences in 22 of 29 transcripts, which meant that most of the transcripts were full-length sequences. One peptide precursor (Allatostatin CC) was known to include a peptide anchor sequence but not a signal sequence. The other 6 transcripts which did not show any signal peptide sequence could be partial sequences without 5' ends.

Table 5.2: The list of predicted neuropeptides in *C. morosus* neuropeptidome.

Abbreviation	**Corresponding Transcript**	**Number of Isoforms**	**Length of Precursor (aa)**	**Signal Peptide Cleavage Site**	**Modification**	**Closest Organism**	**Closest Peptide Code**	**E Value**
AKH	TRINITY_DN28591_c1_g1_i1	4	99	49-50		*Coptotermes gestroi*	AML80828.1	2.00E-12
AST-A	TRINITY_DN34906_c0_g1_i2	3	359	25-26	amidation	*Periplaneta americana*	CAA62500.1	1.00E-53
AST-B	TRINITY_DN36424_c0_g1_i7	7	243	22-23		*Locusta migratoria*	AKN21242.1	3.00E-42
AST-C	TRINITY_DN20153_c0_g1_i1	1	97	30-31	amidation, disulfide bridge	*Athalia rosae*	XP012268507	3.00E-38
AST-CC	TRINITY_DN18536_c0_g1_i1	1	60	peptide anchor	disulfide bridges	*Zootermopsis nevadensis*	KDR09562.1	3.00E-24

Table 5.2: The list of predicted neuropeptides in *C. morosus* neuropeptidome (cont.).

Abbreviation	**Corresponding Transcript**	**Number of Isoforms**	**Length of Precursor (aa)**	**Signal Peptide Cleavage Site**	**Modification**	**Closest Organism**	**Closest Peptide Code**	**E Value**
AT	TRINITY_DN23606_c0_g2_i1	1	>45	nd	amidation	*Schistocerca gregaria*	AKC92815.1	1.00E-15
	TRINITY_DN23606_c0_g1_i1	1	>45	nd	amidation	*Schistocerca gregaria*	AKC92815.1	1.00E-15
	TRINITY_DN67282_c0_g1_i1	1	75	26-27	amidation	*Schistocerca gregaria*	AKC92815.1	3.00E-07
BURSA	TRINITY_DN31293_c0_g1_i1	5	171	47-48		*Blattella germanica*	CUT08823.1	6.00E-72
BURSB	TRINITY_DN1928_c0_g1_i1	1	140	24-25	disulfide bridges	*Z. nevadensis*	KDR13885.1	3.00E-64
CCAP	TRINITY_DN26064_c0_g1_i1	2	153	26-27		*Z. nevadensis*	KDR08645.1	9.00E-47

Table 5.2: The list of predicted neuropeptides in *C. morosus* neuropeptidome (cont.).

Abbreviation	**Corresponding Transcript**	**Number of Isoforms**	**Length of Precursor (aa)**	**Signal Peptide Cleavage Site**	**Modification**	**Closest Organism**	**Closest Peptide Code**	**E Value**
ITP	TRINITY_DN30216_c0_g1_i1	1	118	24-25	disulfide bridges	*Halyomorpha halys*	XP014274477	3.00E-50
DH31	TRINITY_DN28631_c0_g1_i1	1	105	21-22		*Nilaparvata lugens*	AFW19797.1	1.00E-35
DH44	TRINITY_DN31821_c0_g1_i1	1	168	21-22		*Z. nevadensis*	KDR14744.1	2.00E-32
ILP	TRINITY_DN23701_c1_g1_i1	1	131	18-19	two chains, disulfide bridges	*Nilaparvata lugens*	AFW19800.1	5.00E-06
	TRINITY_DN28412_c0_g1_i1	2	119	17-18	two chains, disulfide bridges	*Limulus polyphemus*	XP013774489	2.00E-07

Table 5.2: The list of predicted neuropeptides in *C. morosus* neuropeptidome (cont.).

Abbreviation	**Corresponding Transcript**	**Number of Isoforms**	**Length of Precursor (aa)**	**Signal Peptide Cleavage Site**	**Modification**	**Closest Organism**	**Closest Peptide Code**	**E Value**
ILP	TRINITY_DN29225_c0_g1_i1	1	128	22-23	two chains, disulfide bridges	*Z. nevadensis*	KDR17011.1	2.00E-23
	TRINITY_DN17024_c0_g1_i1	1	>89	nd	two chains, disulfide bridges	*Ceratitis capitata*	XP012159442	5.00E-08
	TRINITY_DN71378_c0_g1_i1	1	>95	nd	two chains, disulfide bridges	*Anoplophora glabripennis*	XP018572613	2.00E-09
MS	TRINITY_DN30329_c0_g1_i1	2	119	43-44	amidation	*Blattella germanica*	CAF04070.1	6.00E-28

Table 5.2: The list of predicted neuropeptides in *C. morosus* neuropeptidome (cont.).

Abbreviation	**Corresponding Transcript**	**Number of Isoforms**	**Length of Precursor (aa)**	**Signal Peptide Cleavage Site**	**Modification**	**Closest Organism**	**Closest Peptide Code**	**E Value**
NP	TRINITY_DN32007_c0_g1_i1	3	108	22-23	disulfide bridges	*Tribolium castaneum*	XP008193973	3.00E-19
NPF	TRINITY_DN30006_c0_g1_i1	2	97	33-34	amidation	*Coptotermes formosanus*	AGM32387.1	8.00E-21
NPY	TRINITY_DN30388_c0_g1_i1	2	117	27-28	amidation	*Z. nevadensis*	KDR17622.1	2.00E-22
ORC	TRINITY_DN29783_c0_g2_i2	1	174	30-31		*Blattella germanica*	AKR13995.1	3.00E-62
ORC	TRINITY_DN29783_c0_g2_i1	1	232	30-31		*Polistes dominula*	XP015174930	4.00E-20

Table 5.2: The list of predicted neuropeptides in *C. morosus* neuropeptidome (cont.).

Abbreviation	**Corresponding Transcript**	**Number of Isoforms**	**Length of Precursor (aa)**	**Signal Peptide Cleavage Site**	**Modification**	**Closest Organism**	**Closest Peptide Code**	**E Value**
sNPF	TRINITY_DN20373_c0_g1_i1	1	99	25-26		*Nilaparvata lugens*	AFW04602.1	1.00E-22
CNM	TRINITY_DN31439_c2_g1_i1	1	176	55-56	amidation	*Athalia rosae*	XP012268550	3.00E-05
TRN	TRINITY_DN1575_c0_g1_i1	1	>93	nd		*Culex quinquefasciatus*	XP_001844331	4.00E-25
7B2	TRINITY_DN68925_c0_g1_i1	1	>71	nd	disulfide bridges	*Zootermopsis nevadensis*	KDR19503.1	7.00E-28

The precursor proteins were processed and controlled if they have the core motives of the corresponding neuropeptide family. The predicted maturation process of AST-A precursor protein is given in Figure 5.7. Initially, the signal peptide sequences were extracted from the predicted open reading frames via SignalP4.1. Secondly, the mono or dibasic cleavage sites which were processed by proprotein convertases were predicted according to the rules stated by Duckert *et al.* in 2004 [33]. Thirdly, C-terminal basic residues were cleaved by carboxypeptidase E. Finally, if there was an amidation site such as a glycine at C-term, this residue was transformed into an amide. We aligned these processed peptides within the different transcript isoforms, also with the peptides of other species. For AST-A example, even one transcript could be processed into 13 mature peptides which have the common C-terminal core motif of FGLamide (Figure 5.8).

The same search was performed for the AST-C peptide which is the original ligand of CamAlstR-C receptor. Only one transcript (TRINITY_DN20153_c0_g1_i1) was obtained for this peptide precursor. After cleavage of the signal peptide sequence 67 residues were left. There was a dibasic cleavage site for proprotein convertase and this lead to an 18-aa peptide at the C-term of the propeptide. Carboxypeptidase E cleaved the terminal Lysine residue and the C-terminal Glycine was amidated. The mature peptide was predicted to be KRSYWKQCAFNAVSCFamide with a disulfide bridge between C8 and C15 (Figure 5.9).

Another example for processing of insulin-like peptide (ILP) precursor sequence is given in Figure 5.10. All predicted ORFs were checked for the presence of disulfide bridges in DIANNA server. Insulin is one of the most known peptides for its maturation process. Its N- and C-terminal regions constitute A and B chains, which fold onto each other with the help of three disulfide bridges. After its packaging and when it's in the secretory vesicles, the loop residues between these chains are cleaved by proprotein convertases. At the last step, carboxypeptidase E cleaves the basic lysine on the chain B. As given in the Figure 5.10, insect ILPs are processed with the same pattern. We have identified 5 different transcripts and one of them has 2 isoforms, as coding for ILPs. A and B chains of these ORFs were aligned and given in Figure 5.11. Positions

Figure 5.7: Processing of AST-A precursor peptide from TRINITY_DN34906_c0_g1_i2. The predicted longest ORF of the transcript is given at the beginning. Blue residues are signal peptide sequences. Putative mature peptides are given in purple. Processing enzymes are given next to yellow arrows.

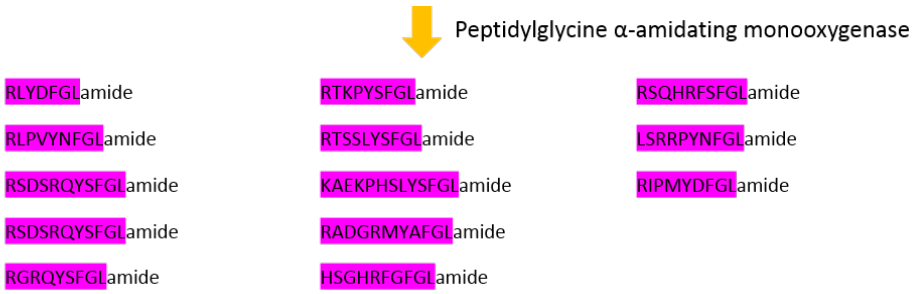


Figure 5.7 (cont.): Processing of AST-A precursor peptide from TRINITY_DN34906_c0_g1_i2. The predicted longest ORF of the transcript is given at the beginning. Blue residues are signal peptide sequences. Putative mature peptides are given in purple. Processing enzymes are given next to yellow arrows.

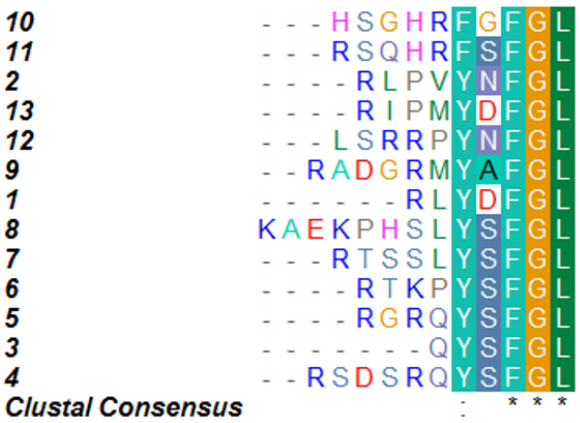


Figure 5.8: Multiple alignment of putative mature AST-A peptides produced from transcript TRINITY_DN34906_c0_g1_i2. Peptides are numbered in their order on the transcript.

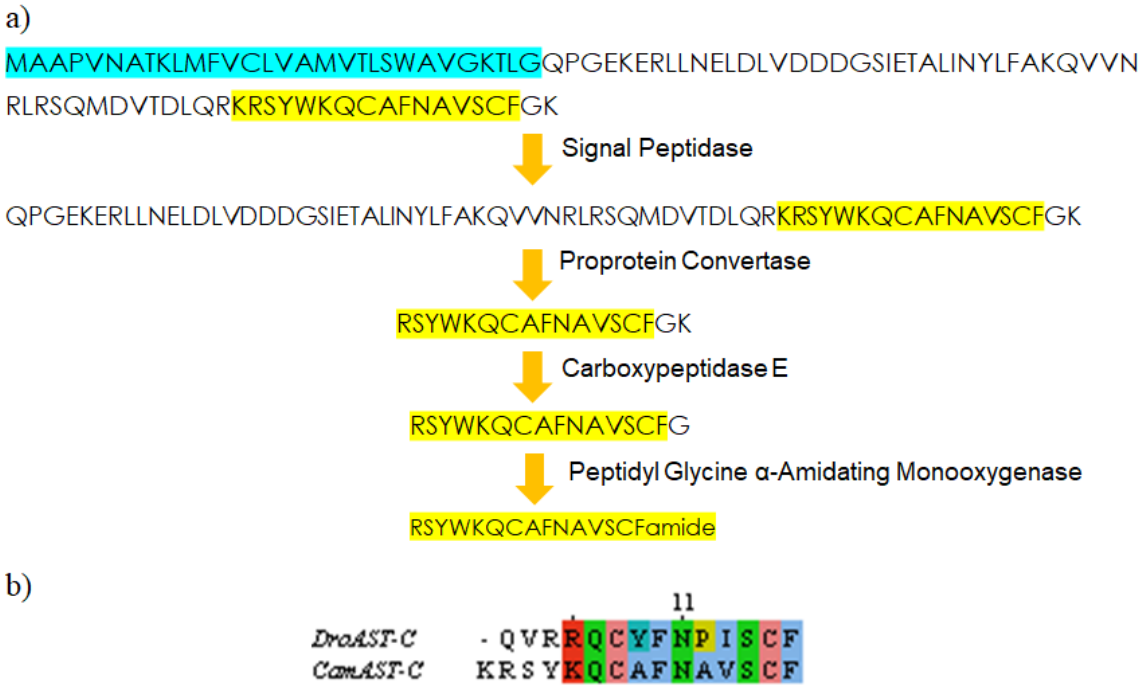


Figure 5.9: Processing of AST-C pre-prohormone.

of 4 Cysteine residues on chain A and 2 Cysteine residues on chain B were conserved within different ORFs.

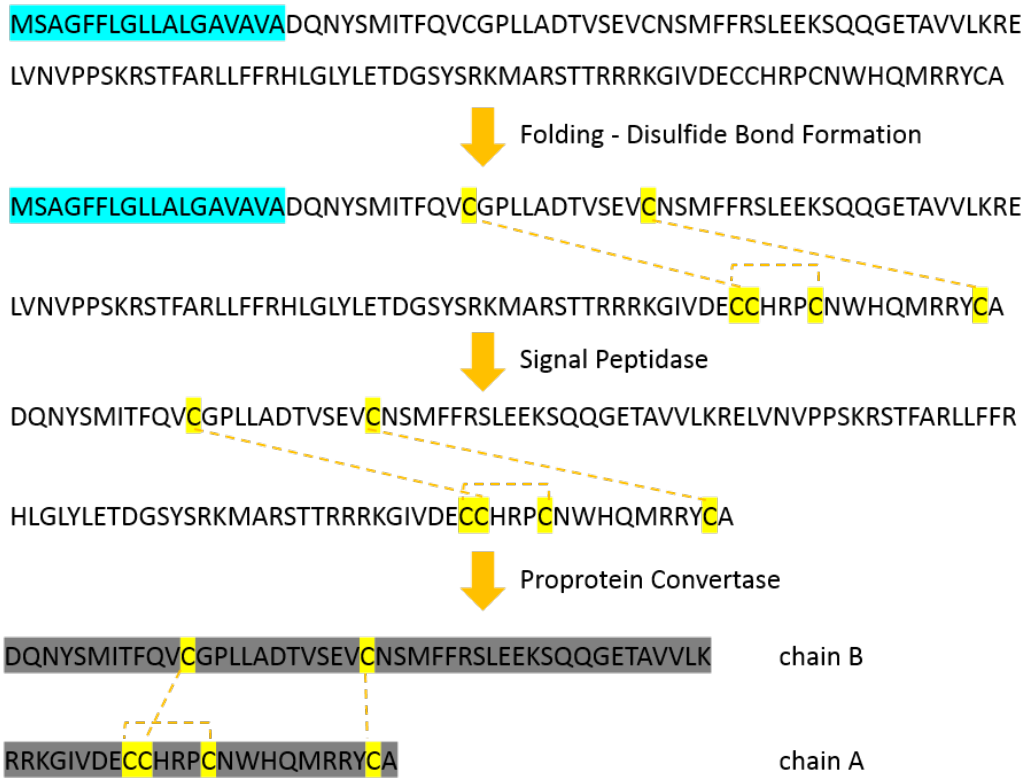


Figure 5.10: Processing of ILP precursor peptide from TRINITY_DN23701_c1_g1_i1. The predicted longest ORF of the transcript is given at the beginning. Blue residues are signal peptide sequences. Disulfide bridges are given in yellow lines. Two mature chains are shown in grey. Processing enzymes are given next to yellow arrows.

Figure 5.11: Multiple alignment of putative mature ILP peptides produced from four different transcripts (TRINITY_DN23701_c1_g1_i1, TRINITY_DN28412_c0_g1_i1, TRINITY_DN29225_c0_g1_i1 and TRINITY_DN71378_c0_g1_i1). A-chain (a) and B-chain (b) alignments were given separately.

5.4. GPCRome of *C. morosus*

The results of GPCRome prediction exhibited 430 putative GPCR transcripts (Figure 5.12a). Within these transcripts, 150 of them were giving highly significant (E≤0.01) similarities in blast search (Figure 5.12b). And 43 of 141 contained full length GPCR structures with 7TM helices (Figure 5.12c). Twenty nine of these were classified in Class A, 10 in Class B1, 2 in Class B2 and 2 in Class C GPCRs. Similarity search showed that one of the 141 highly significant GPCR transcripts was uncharacterized with 3 TM helices. No full length frizzled and Taste-2 receptors could be obtained. Still, 1 Taste-2 GPCR and 3 frizzled receptors were detected from partial transcripts. Types of GPCRs that are expressed in adult *C. morosus* body can be seen in Table 5.4. No steroid and hydroxycarboxylic/nicotinic acid receptors could be detected in the transcriptome, as expected for the arthropods.

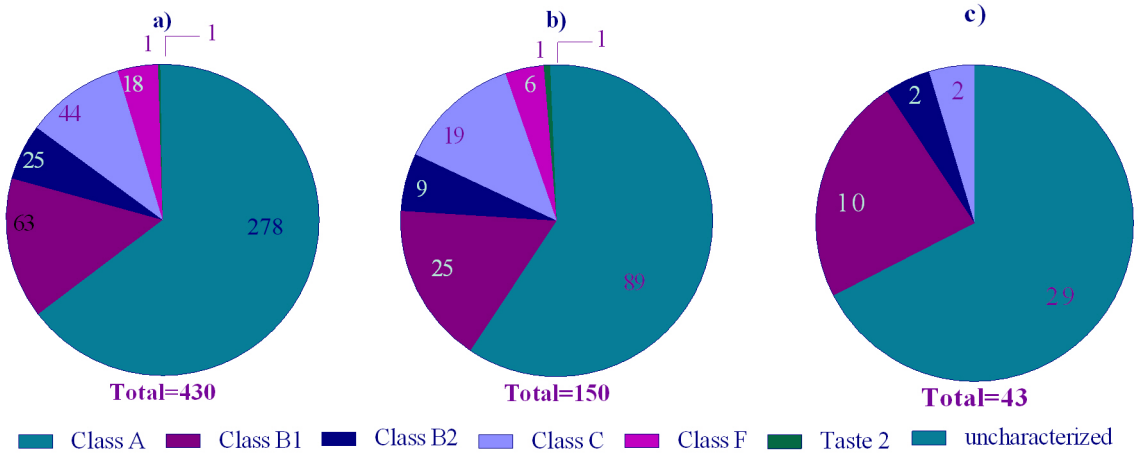


Figure 5.12: Classification of GPCRs that were predicted. a) All of the transcripts giving GPCR hits in blast search. b) The transcripts that yield highly significant (E≤0.01) GPCR hits in blast search. c) The transcripts that yield highly significant GPCR hits in blast search and contain at least 7 helices in their ORFs.

Table 5.3: The types of GPCRs that are obtained from transcriptome of adult *C. morosus* body, and their classification.

Type Of GPCR	**Subclass**	**Class**
GPCR 143	Orphan	
5-Hydroxytryptamine Receptor	Aminergic Receptors	Class A
Adenosine Receptor	Nucleotide Receptors	Class A
Adipokinetic Hormone Receptor	Peptide Receptor	Class A
Allatostatin A Receptor	Peptide Receptor	Class A
Allatostatin C Receptor	Peptide Receptor	Class A
Alpha Adrenergic Receptor	Aminergic Receptors	Class A
Beta Adrenergic Receptor	Aminergic Receptors	Class A
Bombesin Receptor	Peptide Receptor	Class A

Table 5.3: The types of GPCRs that are obtained from transcriptome of adult *C. morosus* body, and their classification (cont.).

Type Of GPCR	**Subclass**	**Class**
Cardioaccelatory Peptide Receptor	Vasopressin/Oxytocin Receptor	Class A
Cephalotocin Receptor	Vasopressin/Oxytocin Receptor	Class A
Chemokine Receptor	Protein Receptor	Class A
Cholecystokinin Receptor Like	Peptide Receptor	Class A
Dopamine Receptor	Aminergic Receptors	Class A
Endothelin Receptor	Peptide Receptor	Class A
Fmrfamide Receptor	Peptide Receptor	Class A
Follicle-Stimulating Hormone Receptor	Peptide Receptor	Class A
Free Fatty Acid Receptor	Lipid Receptors	Class A
Glucose-Dependent Insulinotropic Receptor	Cannabinoid Receptor	Class A
Gonadotropin-Releasing Hormone II Receptor	Peptide Receptor	Class A
Histamine Receptor	Aminergic Receptors	Class A
Inotocin Receptor	Vasopressin/Oxytocin Receptor	Class A
Lutropin-Choriogonadotropic Hormone Receptor	Protein Receptor	Class A
Melanopsin	Sensory Receptors	Class A
Moody	(GPR84)	Class A
Muscarinic Acetylcholine Receptor	Aminergic Receptors	Class A
Neuromedin U Receptor	Peptide Receptor	Class A

Table 5.3: The types of GPCRs that are obtained from transcriptome of adult *C. morosus* body, and their classification (cont.).

Type Of GPCR	**Subclass**	**Class**
Neuropeptide A10/Sex Peptide Receptor	Peptide Receptor	Class A
Neuropeptide A32 Receptor	Peptide Receptor	Class A
Neuropeptide A6a	Peptide Receptor	Class A
Neuropeptide Capa Receptor	Peptide Receptor	Class A
Neuropeptide Cchamide-1 Receptor	Peptide Receptor	Class A
Neuropeptide F Receptor	Peptide Receptor	Class A
Neuropeptide FF Receptor	Peptide Receptor	Class A
Neuropeptide Receptor	Peptide Receptor	Class A
Neuropeptide Receptor A27	Peptide Receptor	Class A
Neuropeptide Sifamide Receptor	Peptide Receptor	Class A
Neuropeptide Y Receptor	Peptide Receptor	Class A
Octopamine Or Capa Receptor	Adrenoreceptor/Vasopressin	Class A
Octopamine Receptor	Adrenoreceptors	Class A
Odorant	Aminergic Receptors	Class A
Odorant Receptor	Odorant Receptor	Class A
Odorant Receptor 4	Odorant Receptor	Class A
Odorant Receptor 40	Sensory Receptor	Class A
Odorant Receptor 83a	Sensory Receptor	Class A
Opsin	Sensory Receptors	Class A
Prolactin-Releasing Peptide Receptor	Peptide Receptor	Class A
Relaxin Receptor	Peptide Receptor	Class A
Rfamide Receptor	Peptide Receptor	Class A

Table 5.3: The types of GPCRs that are obtained from transcriptome of adult *C. morosus* body, and their classification (cont.).

Type Of GPCR	**Subclass**	**Class**
Rhodopsin	Sensory Receptors	Class A
Ryamide Receptor	Neuropeptide Y Receptor	Class A
Sex Peptide Receptor	Peptide Receptor	Class A
Sifamide Receptor	Peptide Receptor	Class A
Tachykinin-Like Peptides Receptor	Peptide Receptor	Class A
Thyrotropin Receptor	Protein Receptor	Class A
Trace Amine Associated Receptor	Aminergic Receptors	Class A
Tyramine Receptor	Adrenoreceptors	Class A
Vasopressin/Oxytocin Receptor	Peptide Receptor	Class A
Calcitonin Receptor	Peptide Receptor	Class B1
Diuretic Hormone Receptor	Peptide Receptor	Class B1
Mth Like	Methuselah-Like	Class B1
PDF Receptor	VIP And PACAP Receptor	Class B1
Pigment Dispersing Factor Receptor	VIP And PACAP Receptor	Class B1
Adhesion GPCR G2	Adhesion Receptor	Class B2
Adhesion GPCR A3	Adhesion Receptor	Class B2
GABA-B Receptor	Amino Acid Receptor	Class C
Gustatory Receptor	Sensory Receptor	Class C
Gustatory Receptor 2	Sensory Receptor	Class C
Gustatory Receptor 28b	Sensory Receptor	Class C
Gustatory Receptor 43a	Sensory Receptor	Class C
Gustatory Receptor 64e	Sensory Receptor	Class C
Gustatory Receptor 64f	Sensory Receptor	Class C

Table 5.3: The types of GPCRs that are obtained from transcriptome of adult *C. morosus* body, and their classification (cont.).

Type Of GPCR	**Subclass**	**Class**
Metabotropic Glutamate Receptor	Amino Acid Receptor	Class C
Frizzled	Frizzled Receptors	Class F
Frizzled-10	Frizzled Receptors	Class F
Gustatory Receptor 28a	Sensory Receptor	Taste 2

The number of helices obtained from the TMHMM analysis of predicted ORFs was given in Table 5.4. The biggest number of helices was 14, which contributed to the membrane spanning transporter proteins. And the least number of helices was 1, most of which came from the signal sequence of pre-propeptides. In order not to miss any partial transcript, blastp search was performed on all of the putative sequences. The result of the number of helices within the expressed GPCRome of adult *C. morosus* was given in Table 5.4.

Table 5.4: Numbers of transcripts that include the corresponding numbers of helices. (The non-GPCR helices were excluded.)

Number of Helices	**Number of Transcripts**
10	9
9	7
8	15
7	18
6	8
5	20
4	24
3	39
2	97
1	197

Table 5.5: The types of GPCRs chosen for tissue specific expression analysis and their transcripts. In presence of multiple isoforms, the primers were designed to amplify all of them.

Type of Receptor	**Class of GPCR**	**Transcript Code**	**# of Isoforms**
Octopamine Receptor	Class A	TRINITY_DN30951_c0_g1	1
Tyramine Receptor 2-like	Class A	TRINITY_DN31442_c1_g1	2
Adipokinetic Hormone Receptor	Class A	TRINITY_DN36998_c1_g1	4
Allatostatin A Receptor	Class A	TRINITY_DN62595_c0_g1	1
Allatostatin C Receptor	Class A	TRINITY_DN42122_c0_g1	1
Inotocin Receptor	Class A	TRINITY_DN36849_c0_g1	5
Neuropeptide Y Receptor	Class A	TRINITY_DN21880_c0_g1	1
Sex Peptide Receptor	Class A	TRINITY_DN54154_c0_g1	1
Cholecystokinin Receptor like	Class A	TRINITY_DN35009_c0_g2	3
Calcitonin Gene-Related Peptide Type 1 Receptor	Class B1	TRINITY_DN35728_c0_g1	1
Diuretic Hormone Receptor	Class B1	TRINITY_DN29760_c0_g1	1
Adhesion GPCR G2-like	Class B2	TRINITY_DN19522_c0_g1	1

Table 5.5: The types of GPCRs chosen for tissue specific expression analysis and their transcripts. In presence of multiple isoforms, the primers were designed to amplify all of them (cont.).

Type of Receptor	**Class of GPCR**	**Transcript Code**	**# of Isoforms**
Orphan GPCR	Uncharacterized	TRINITY_DN65134_c0_g1	1

Some of these GPCRs have more than one isoforms having some amino acid sequence variations (examples in Table 5.4). One of the most variable ones is the glucose-dependent insulinotropic receptor. Four different receptor sequences have deletions in different parts of the receptors, but these variations are confined to the N terminal or C terminal loops. On the other hand, inotocin receptors show only one amino acid differences in their sequences. The GPCRs in Table 5.4 were chosen for tissue specific expression analysis. We tried to choose at least one GPCR from each class. The other criteria were to include the full-lenght transcripts or the GPCRs that are the focus of our previous studies. Therefore we included the GPCRs such as adipokinetic hormone receptor (AKHR, Class A) inotocin receptor (Class A), CCHamide receptor (Class A), octopamine receptor (OctR, Class A), Tyramine receptor (TyrR, Class A), calcitonin receptor (Class B1), diuretic hormone receptor (DHR, Class B1) and gustatory receptor for sugar taste (Class C). Also, other receptors which didn't show all 7TM domains were chosen to help further studies; such as sex peptide receptor (SPR, Class A), Allatostatin A and C receptors (AlstR-A and AlstR-C, Class A), Neuropeptide Y receptor (NYR, Class A), an uncharacterized receptor (orphan GPCR), adhesion GPCR (Class B2) and Frizzled 10 (Class F).

5.4.1. Tissue expression profiles of the predicted GPCR transcripts

The information about the anatomy of *C. morosus* was limited to the gastrointestinal system. In our study, we could discover the major organs of this organism.

Figure 5.13 shows the anatomy of adult *C. morosus* female. Since the animal reproduce parthenogenetically and females can reproduce from unfertilized eggs, the male animals are rarely seen and we didn't have any males. Therefore, we couldn't discover the anatomy of male reproductive system organs. In our tissue expression profile study, we included the organs illustrated in Figure 5.13.

During the dissection, the brain was collected together with the neuroendocrine glands *Corpora allata* and *Corpora cardiaca*. In addition, ovary samples also included the mature eggs because the samples were adult females which were full of eggs. Therefore, ovary samples could show not only the gonad genes but also developmental genes expressed in the eggs.

Before performing expression analysis the RNA samples were evaluated for integrity. The samples did not show sign of degradation (Figure 5.14). The one-band pattern is a common behavior of insect 28S rRNA. In addition, three housekeeping genes were checked (Figure 5.15). Within 3 genes, GAPDH was chosen as the least varying within different tissue samples.

Figure 5.16 shows the fold expression levels of different GPCRs with regard to GAPDH in ovary samples. CCHamide Receptor and Frizzled10 primers did not give a Cp value less than 30 in any of the tissues, so they were excluded from the analysis. The PCR efficiency of the gustatory receptor for sugar taste 43a primers could not be calculated, because its standard curve could not be constructed due to experimental difficulty. Therefore, this receptor was not included in qPCR experiments (Appendix D). Among the others, some GPCRs showed tissue-specific expression profiles, such as CamInoR, CamCalR, CamTyr2R, and CamAKHR. CamInoR was significantly expressed in gastric cecea while CamTyr2R was in the aorta, CamCalR in the fat body and CamAKHR in the fat body as well as the ganglia. CamSPR expression in brain, CA and CC sample was significantly higher than in any of the other organs. When compared to ovary levels, it was highly expressed in ganglia, crop, foregut, fat body and aorta. Expression of CamAdgrG2 was higher in ganglia and post-posterior midgut + hindgut than the other organs but the difference was not statistically significant.

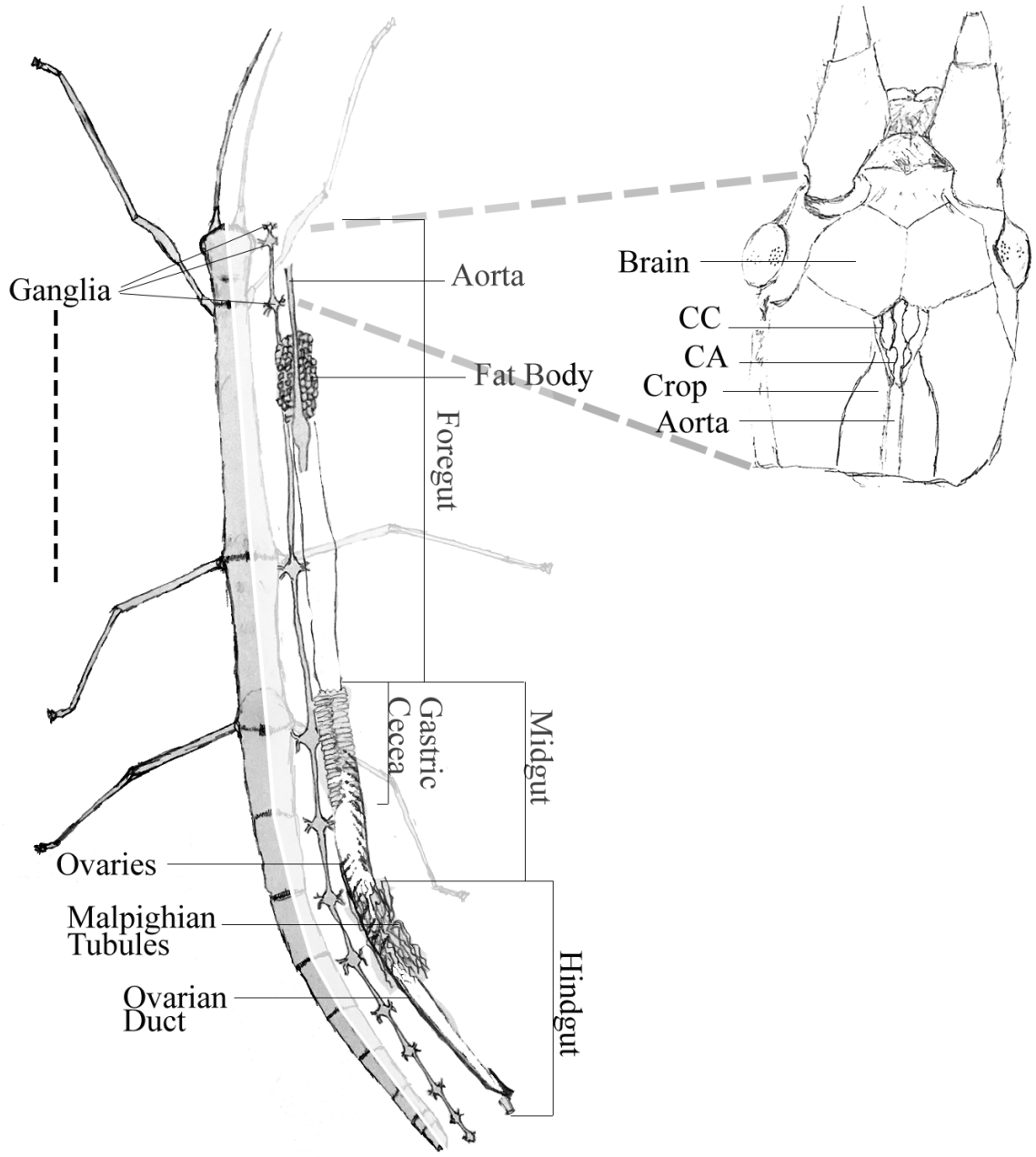


Figure 5.13: Anatomy of the female stick insect, *C. morosus*. Only the organs that were included in RNA isolation were illustrated. CC: Corpus cadiacum, CA: Corpora allatum.

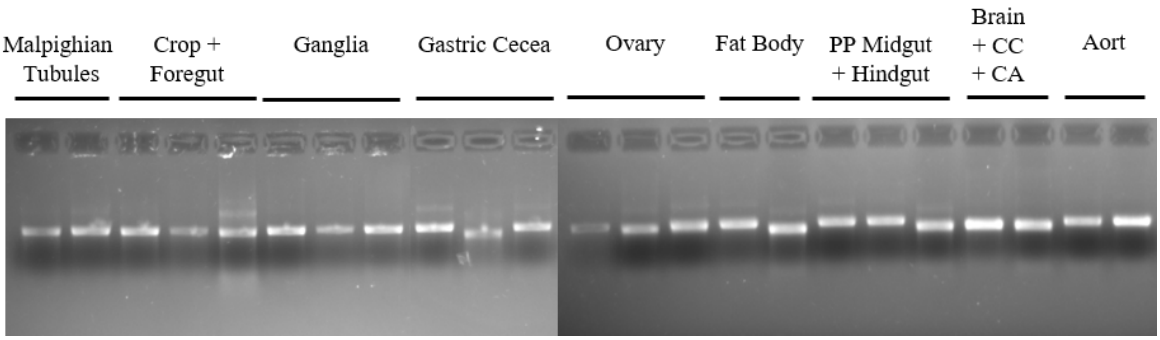


Figure 5.14: RNA samples of tissues were run in MOPS gel electrophoresis in denaturing conditions.

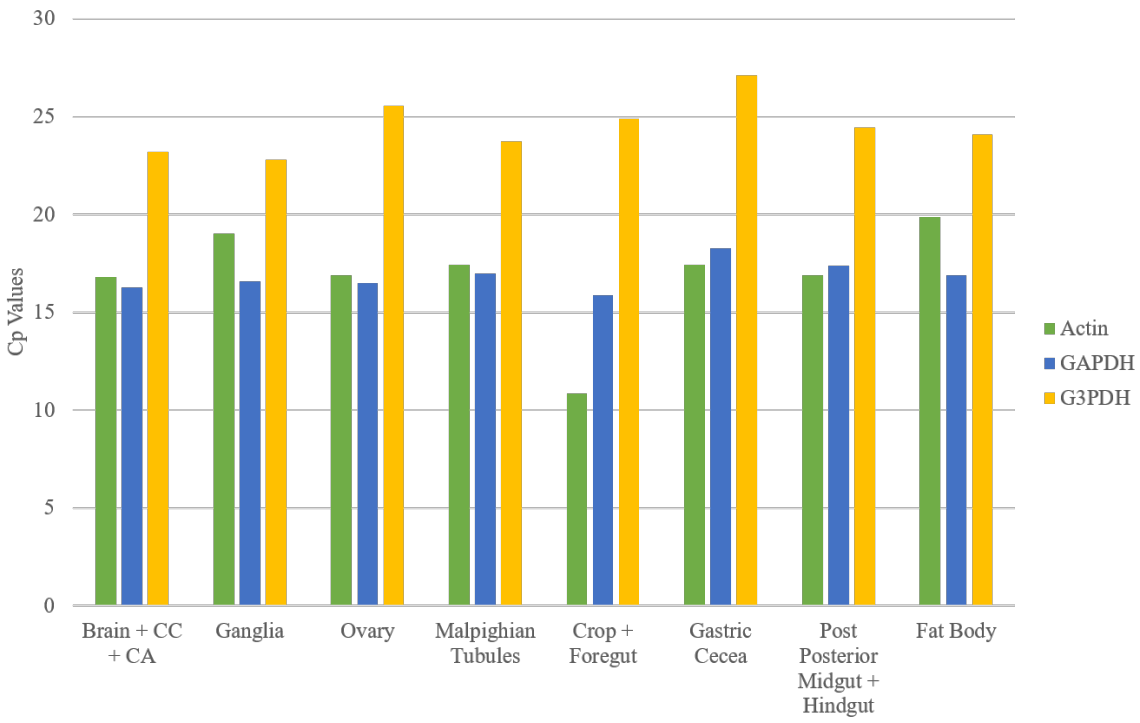


Figure 5.15: Cp values in expression of housekeeping genes.

CamDHR level was higher in Malpighian tubules and post-posterior midgut + hindgut than the other organs but the difference was not significant again. CamCCKR was higher in gastric cecea (including the anterior midgut) but the difference was insignificant. The others, CamNPYR, CamAlstR-A, CamAlstR-C, and CamOctR were more uniformly expressed within the tissues, than the other GPCRs in-analysis. The challenging discovery was the tissue-specific expression of the orphan GPCR that was predicted from the *in silico* analyses. It was significantly expressed the brain, CC, CA and ganglia samples.

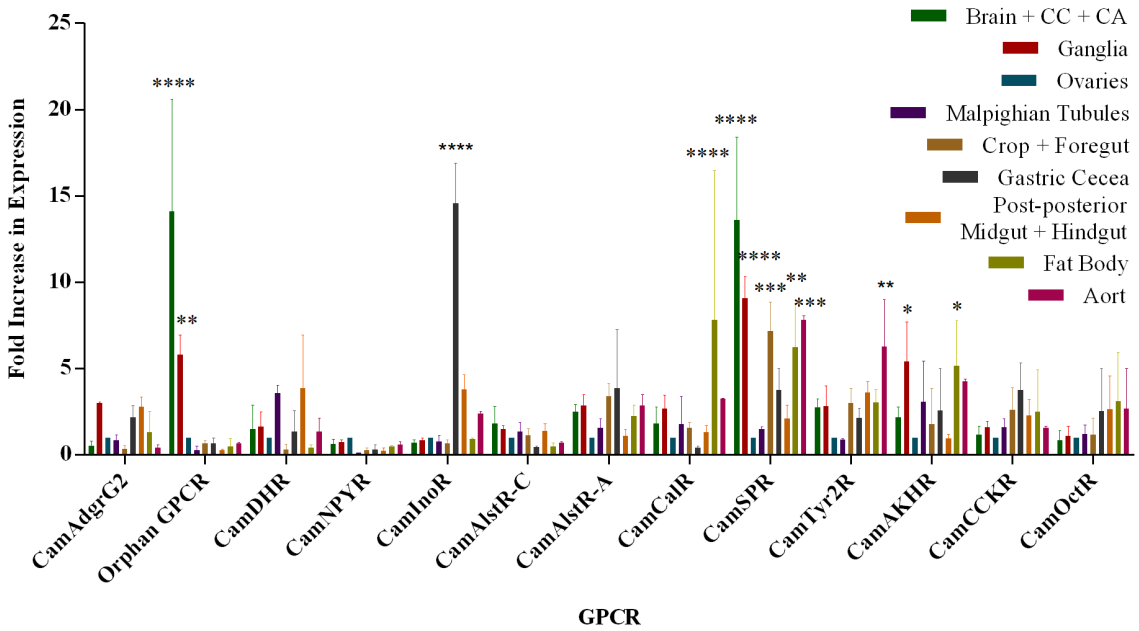


Figure 5.16: Fold difference in expression of GPCR gene relative to GAPDH (=reference) in ovary (=calibrator) via REST Method. The expression levels of GPCRs were compared via two-way ANOVA with Tukey test (* $p \leq 0.05$). n=2 for Malpighian Tubules and Aorta samples, but 3 for the rest of the organs.

Additionally, RT-PCR experiments were performed to confirm expression of some of these GPCRs. In Figure 5.17a, Gustatory receptor for sugar taste 43a showed specific expression in gastric cecea and very low expression in post-posterior midgut and hindgut. This expression profile was also evident in qPCR experiments but the correlation between these patterns was not significant (p=0.069). On the other hand,

brain specific expression of Orphan GPCR was also evident in RT-PCR experiments (Figure 5.17b), and these two types of experiments gave correlated expression patterns (p=0.011).

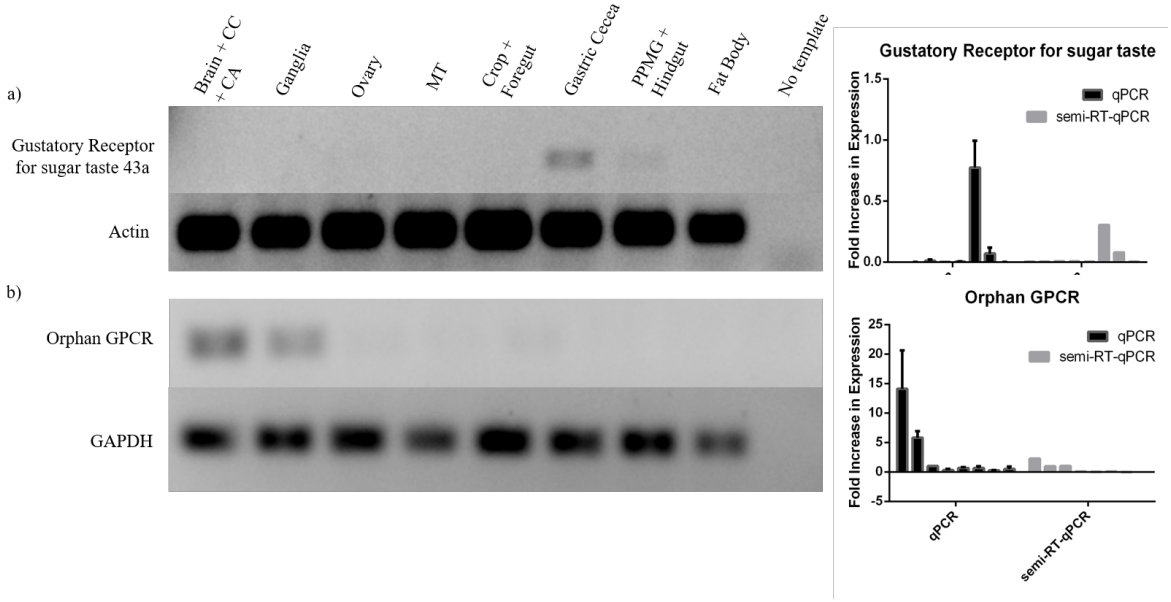


Figure 5.17: Semi-RT-qPCR results of (a) Gustatory receptor for sugar taste 43a and (b) Orphan GPCR. The statistical analysis between the RT-PCR and qPCR graphs were performed in Spearman correlation test. The graphs of Orphan GPCR was found to be correlated ($p<0.05$).

5.5. *In vivo* Effect of AlstR-AST System on Tumor Growth

In the following part of the study, the aim was to discover the effects of AlstR-AST system on proliferation of cells. Therefore, for the functional studies, we planned to expressed this receptor in mammalian cells which do not express it intrinsically. The other goals of this part were to obtain a functional receptor on cell membrane and to analyze its effect on cell viability and cell cycle. The last goal was to observe its effect on tumor xenografts *in vivo*.

5.5.1. Verification of CamAlstR-C Expression in Stable Huh7 Cell Lines

Huh7 cells, which were transfected with CamAlstRC-IRES-EGFP and EGFP-IRES-EGFP vectors, were selected in Geneticin medium for stable transfection. One day after the transfection GFP signal was sufficient to predict over 70% of transfection efficiency (Figure 5.18a). However, during the selection period IRES activity decreased and after one month of selection GFP signal of the cells has decreased down to about 36% in FACS measurements (Figure 5.18b). Reverse Transcription PCR for these cells showed the presence of CamAlstR-C mRNA in AlstR-stable cells and absence of this mRNA in GFP-stable cells (Figure 5.19). However, this expression could not be verified in protein level. In WB membranes, any specific bands for HA-fused-CamAlstRC couldn't be detected (data not shown). Transient transfection of pcDNA3-HA-CamAlstRC was performed into Huh7 cells and WB analysis was repeated for these cells. However, the result was the same (data not shown). The constructs were sent for sequencing and HA epitope was in frame with the gene of interest.

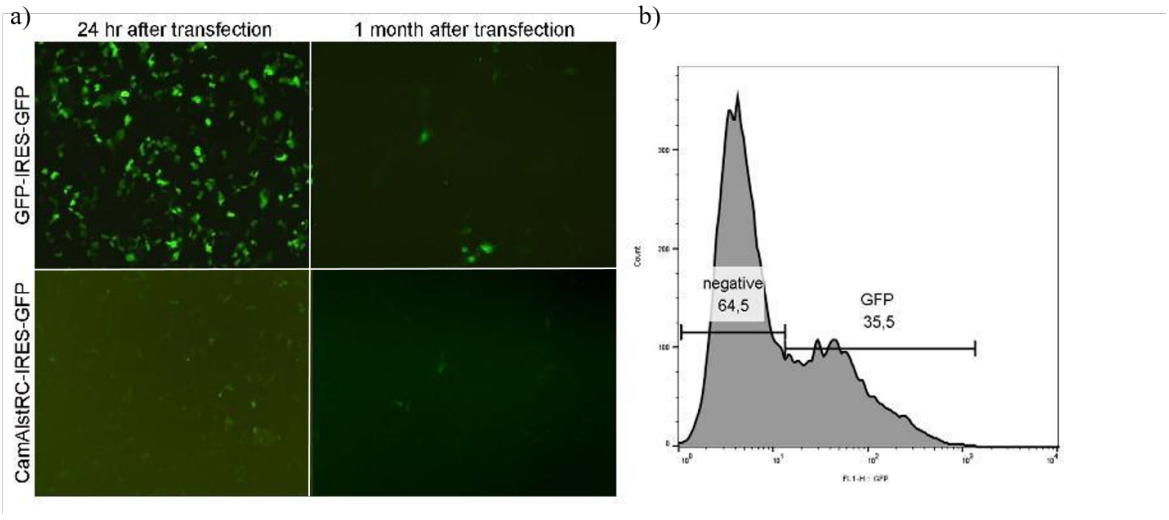


Figure 5.18: Transfection efficiency via (a) fluorescent microscopy and (b) flow cytometry. The images of GFP-IRES-GFP transfected and CamAlstRC-IRES-GFP transfected Huh7 cells were taken at different times. The flow cytometry histogram shows the GFP positive cell population 1 month after transfection.

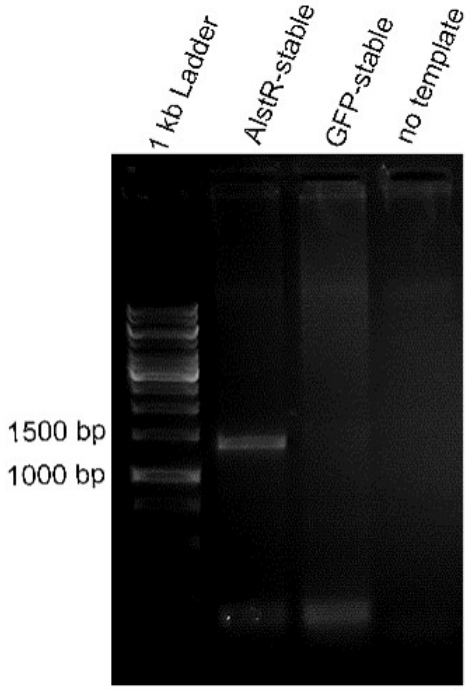


Figure 5.19: RT-PCR for CamAlstRC expression in mRNA level in stable cell lines. 1 kb Ladder, the product of AlstR-stable cell cDNA, GFP-stable cell cDNA and no template PCR reactions were load into the wells, respectively. The predicted product length of CamAlstR-C is 1338 bp.

5.5.2. Effect of CamAlstRC-ASTC System on Proliferation of Hepatocellular Carcinoma Cells

In our preliminary experiments, we have seen that AST-C treatment to Drosophila BG3 cell lines, which are intrinsically AlstR-expressing central nervous system cells, resulted in a response that is similar to insulin deprivation on these cell lines (data not shown). This response was more likely a G2 arrest response. Then we asked if this effect can be replicated on mammalian cancer cells, especially Huh7 cell line. In order to assess cell cycle information, FACS measurements were performed via PI staining. Empty plasmid transfected cells retained in S phase with a frequency of 6.17 % when they were mock treated (Figure 5.20). Their frequency in S phase did not change so much when they were treated with AST-C. However, when they express CamAlstR-C they retained in the S phase more than empty plasmid transfected cells. And when these CamAlstR-C expressing cells were treated with AST-C peptide, the frequency of cells in S phase decreased and shifted towards the G2 phase (52.4 %).

During fixation and PI staining of stable Huh7 cell lines, the clump-forming properties of the cells have changed. As a result, FACS measurements could not give significant and reliable data on stable cell lines. Especially, AlstR-stable cells formed cell clumps very easily and individual cells could not be detected in FACS. Therefore, XTT assays were performed with regard to AST-C treatment on stable cells (Figure 5.21). Increasing amounts of AST-C (such as 10 μM final concentration) lead to a decrease in cell proliferation of AlstR-stable cells. Higher amounts such as 100 μM killed both GFP-stable and AlstR-stable cells in 48 hr. Therefore, we decided to continue with 10 μM concentration during *in vivo* treatments.

5.5.3. *In vivo* Effect of CamAlstRC-ASTC System on Tumor Growth

The growth of both AlstR-tumor and GFP-tumors were very random when they were not treated. In the first three mice (Mouse #1, #2 and #3), two were used for AST-C treatment and another for mock treatment. AST-C treatment inhibited the growth of both AlstR-tumor and GFP-tumor (Figure 5.22) and made its size smaller

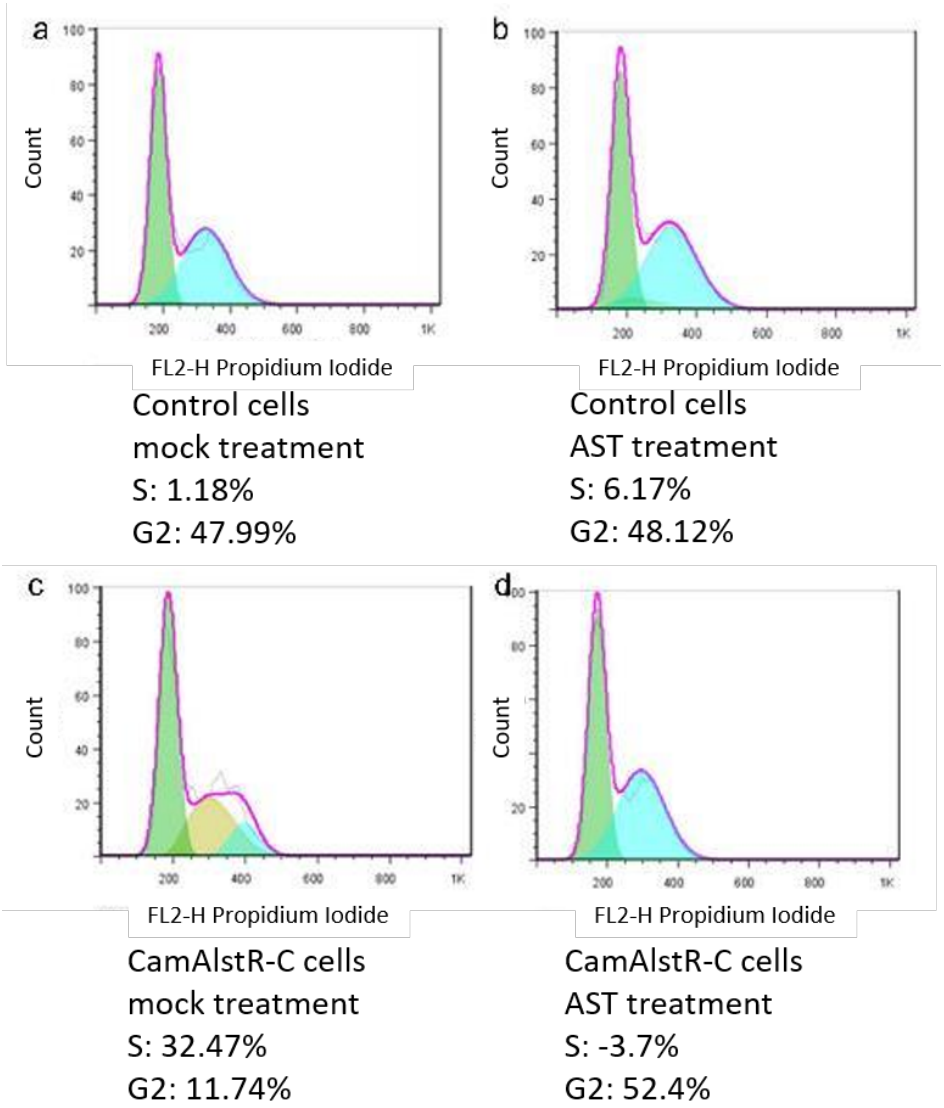


Figure 5.20: The effect of CamAlstRC-ASTC system on cell cycle. CamAlstR-C expressing (c and d) and empty-plasmid transfected (control) cells (a and b) were treated with AST (b and d) or PBS (a nad c) and analyzed in flow cytometry.

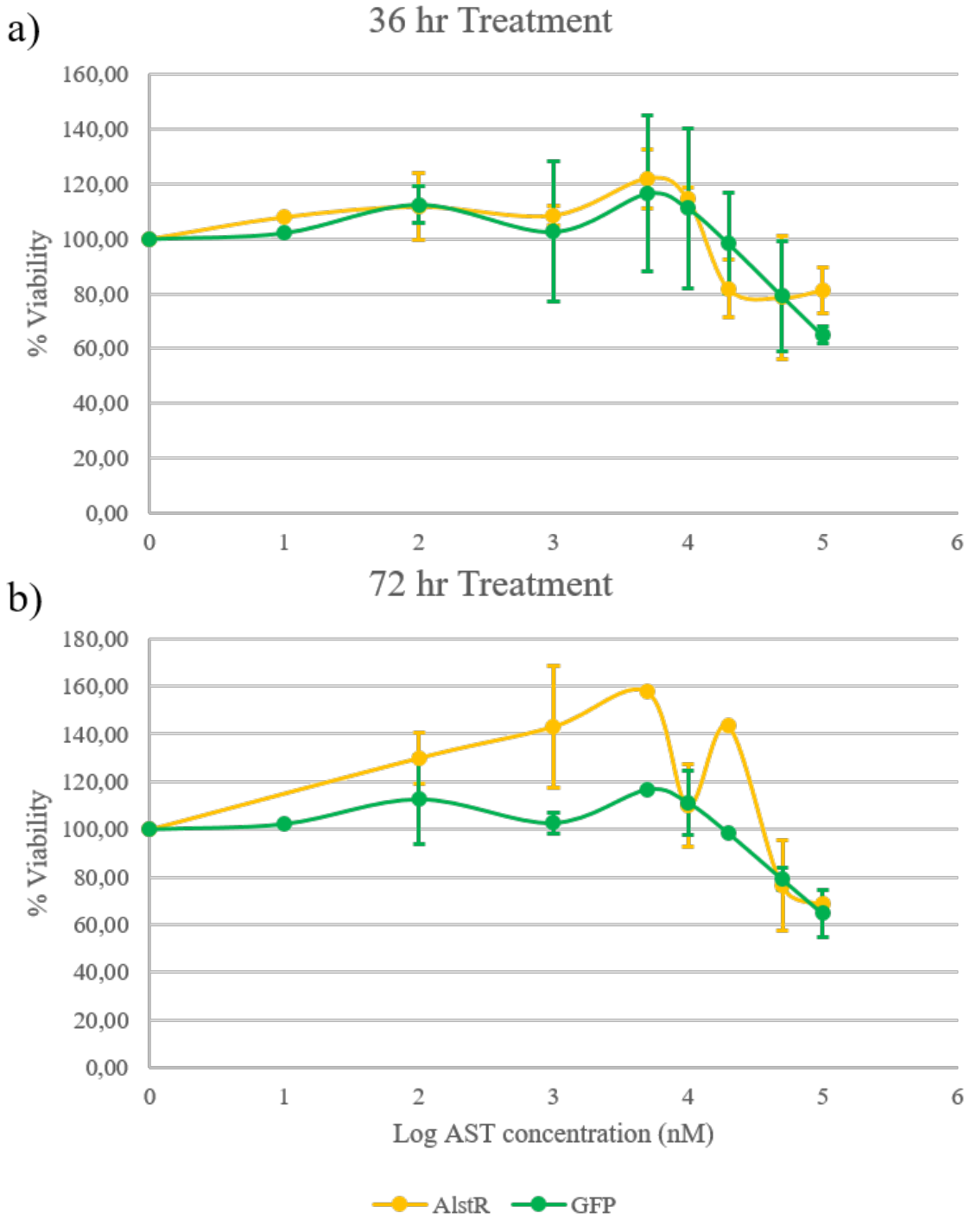


Figure 5.21: The effect of CamAlstRC-ASTC system on proliferation of Huh7 cells. AlstR-stable and GFP-stable cells were treated with various concentrations of AST-C or PBS and analyzed in XTT. 36 hr and 48 hr AST-C treatment results were given in terms of percent viability. The error bars were calculated as S.D.

than the size on the first day of treatment (Figure 5.23 and 5.22) of Mouse #1, but the results were not the same for Mouse #2. In comparison, tumors of saline-treated Mouse #3 maintained growth (Figure 5.24 and 5.25). Another observation of the tumors from Mouse #1 was that the isolated tumors showed no signs of veins/blood when compared to the isolated tumors of Saline-treated Mouse #3.

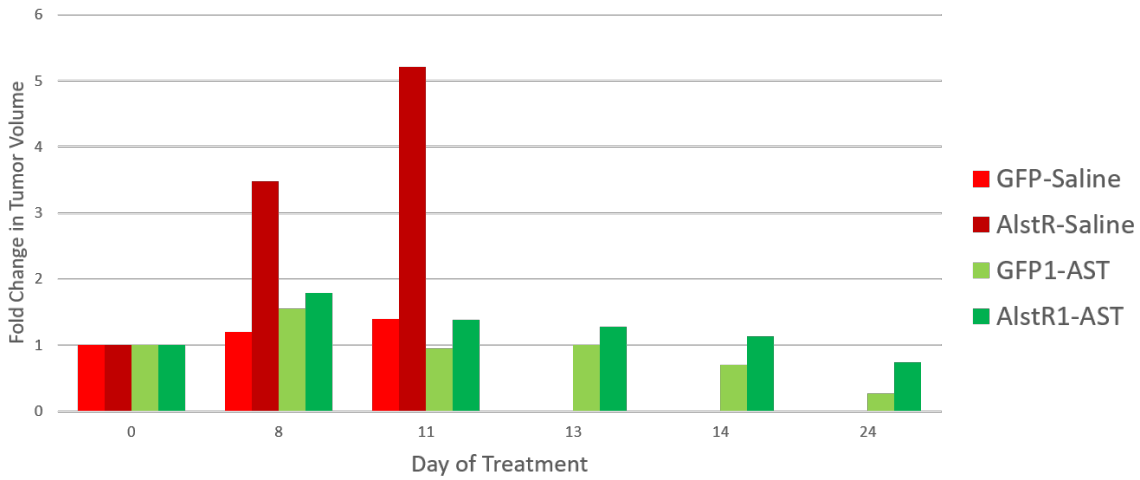


Figure 5.22: Tumor fold changes of Mouse #1 (ASTC treatment – green bars) and #3 (Saline treatment – red bars). Tumor volume of each side of the animals were normalized to 1. (These results are repeated and combined with the other mouse results in Figure 5.24 and 5.25)

Due to the results obtained from Mouse #1, another question came into consideration. AST-C peptide (10 μM)was acting on both AlstR- and GFP-tumors. Another set of xenograft experiment was planned for other 10 mice (from #4 to #13). In order to eliminate the effect of GFP gene (double GFPs) on tumor response, we designed another tumor cell line which expresses only one mCherry gene. In order to eliminate the effect of stable transfection, we also designed another grafting with WT Huh7 cells. GFP emission under the skin of animal could not be detected in live imaging systems. Therefore, mCherry-tumors would also work for comparison of two different volumetric measurement methods, one with a caliper and the other from fluorescent emission (Figure 5.26). The mCherry tumor of Mouse #9 grew in two globular pieces and the Rmax that was taken from the distal ends of the two globes was increasing during the

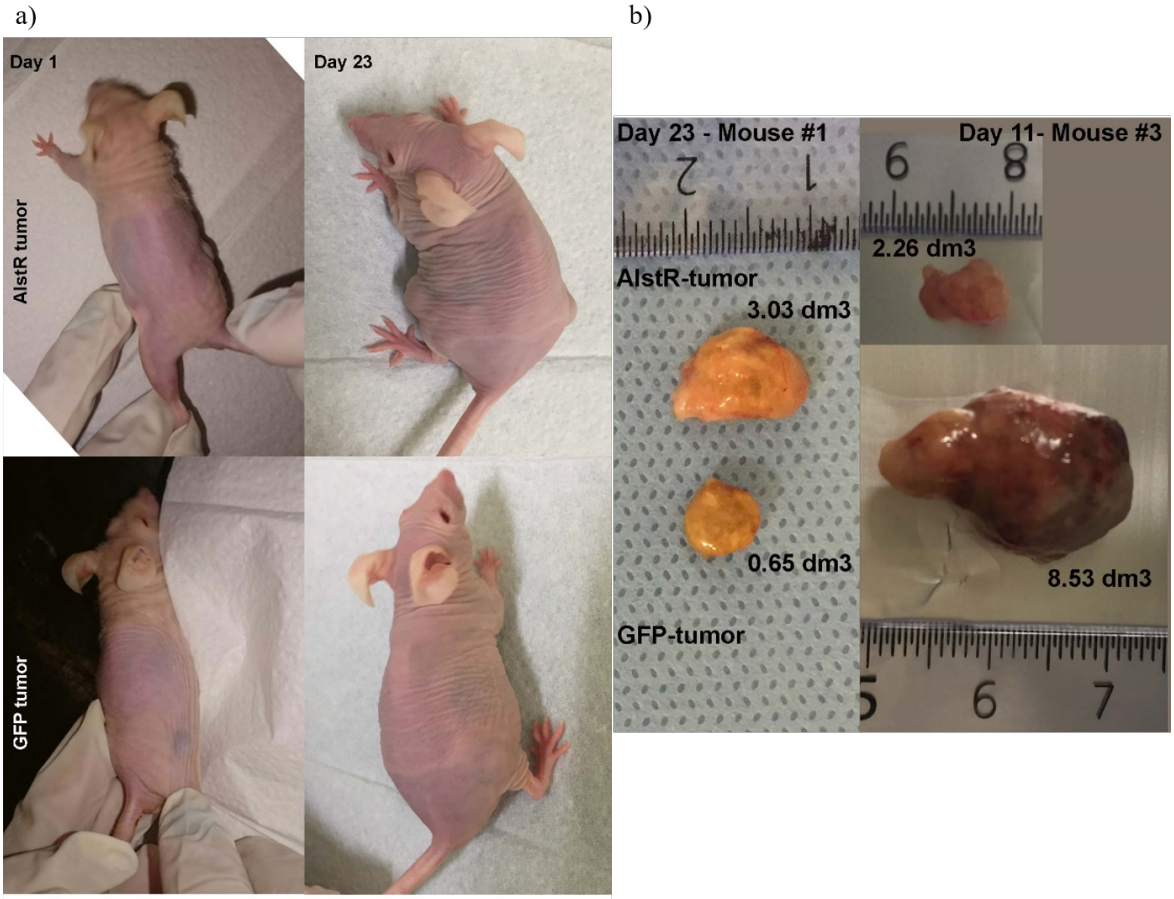


Figure 5.23: Photographs of a) Mouse #1 on 1st and 23rd day of AST-C treatment. AlstR-stable tumor was on the left flank side of the animal, while GFP-stable was on the right. b) Volume measurements of AlstR (top) and GFP (bottom) tumors were shown after isolation (23rd day of AST) and #3 (11th day of Saline).

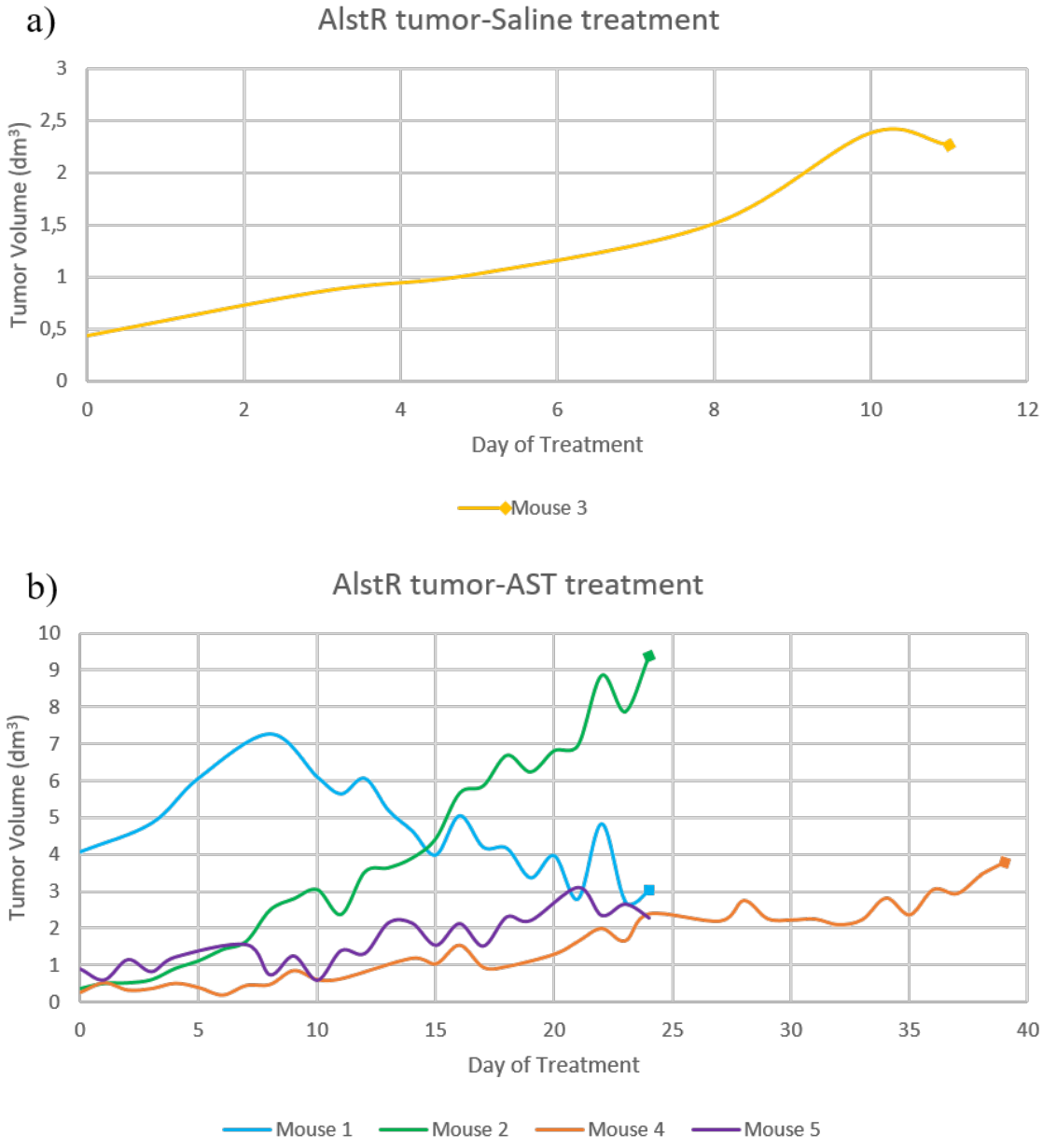


Figure 5.24: AlstR-tumor volumes (dm^3) after AST-C or Saline treatment, of Mouse #1 to #5. A square at the end of curve means that the mouse was sacrificed.

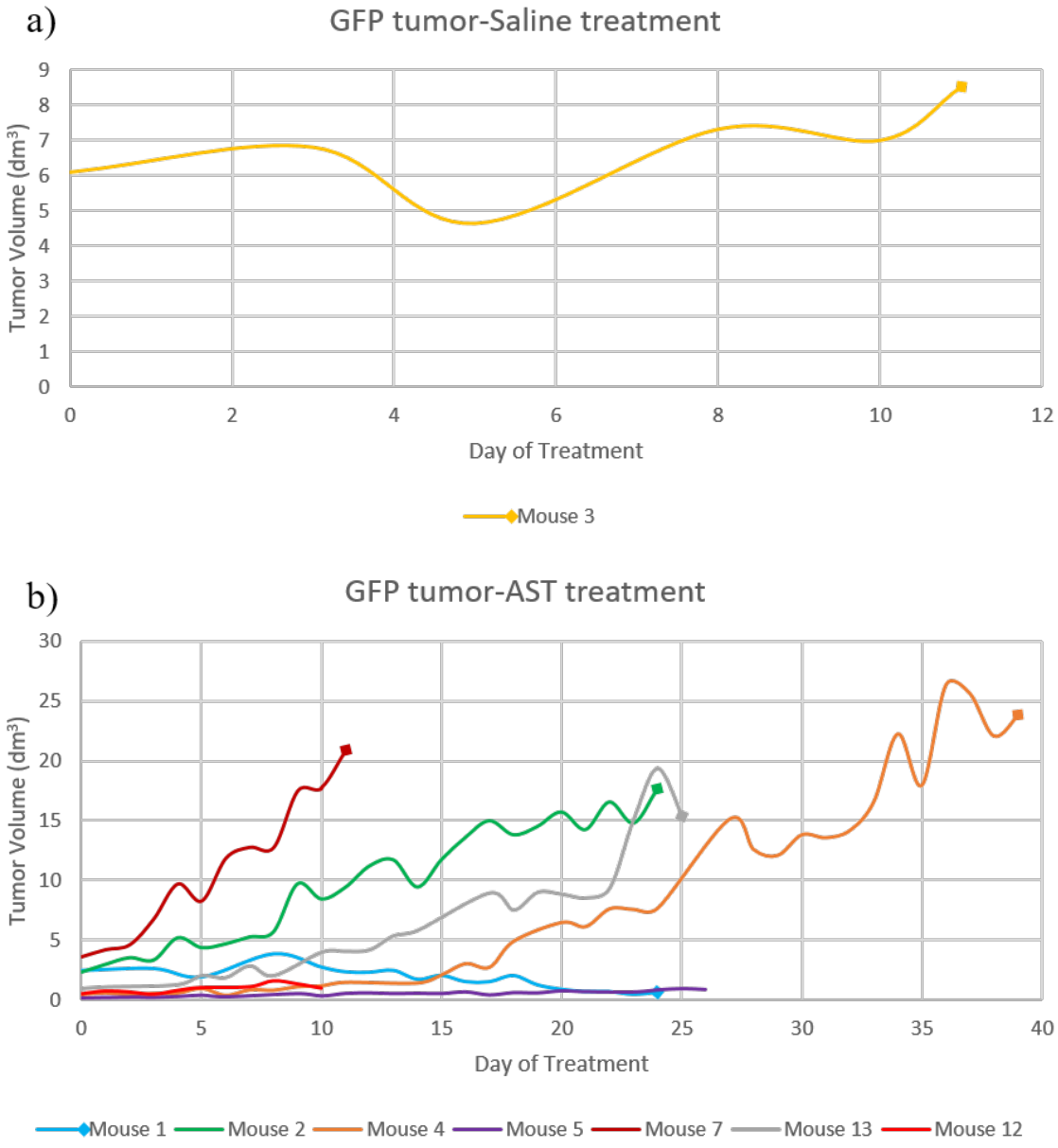
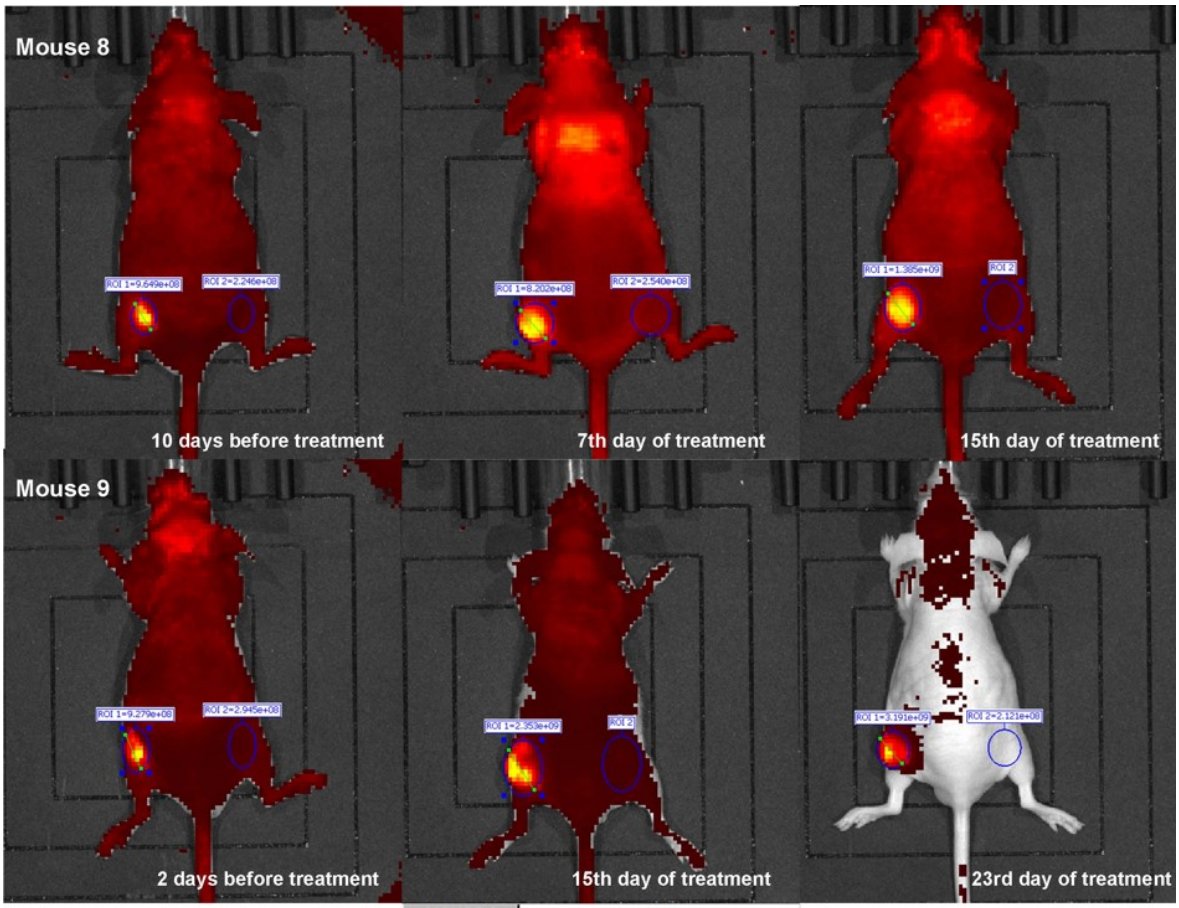


Figure 5.25: GFP-tumor volumes (dm^3) after AST-C or Saline treatment, of Mouse #1 to #5 and #7, #12 and #13. A square at the end of curve means that the mouse was sacrificed.

first 7 days of AST treatment (Figure 5.27). Intratumoral treatment was begun on the 11th day. From the 7th day, the smaller globe began shrinking and almost disappeared in the fluorescent images on the 23rd day of treatment (Figure 5.26). However, the total volume of tumor measured by caliper seemed to be stabilized in Figure 5.27. This response on the 7th day of treatment was similar to the responses obtained in Mouse #1. In contrast, the mCherry tumor of Mouse #8 did not show any stabilization or shrinkage.



MOUSE 8	Left		
	10 days before	7th day of treatment	15th day of treatment
Total Radiant Efficiency [p/s] / [μW/cm²]	7.40E+08	5.66E+08	1.11E+09
Area [cm²]	4.20E-01	6.77E-01	8.01E-01

MOUSE 9	Left		
	2 days before	15th day of treatment	23rd day of treatment
Total Radiant Efficiency [p/s] / [μW/cm²]	6.33E+08	1.93E+09	2.98E+09
Area [cm²]	5.53E-01	1.03E+00	5.72E-01

Figure 5.26: mCherry-tumor images of Mouse #8 and #9 under IVIS system.

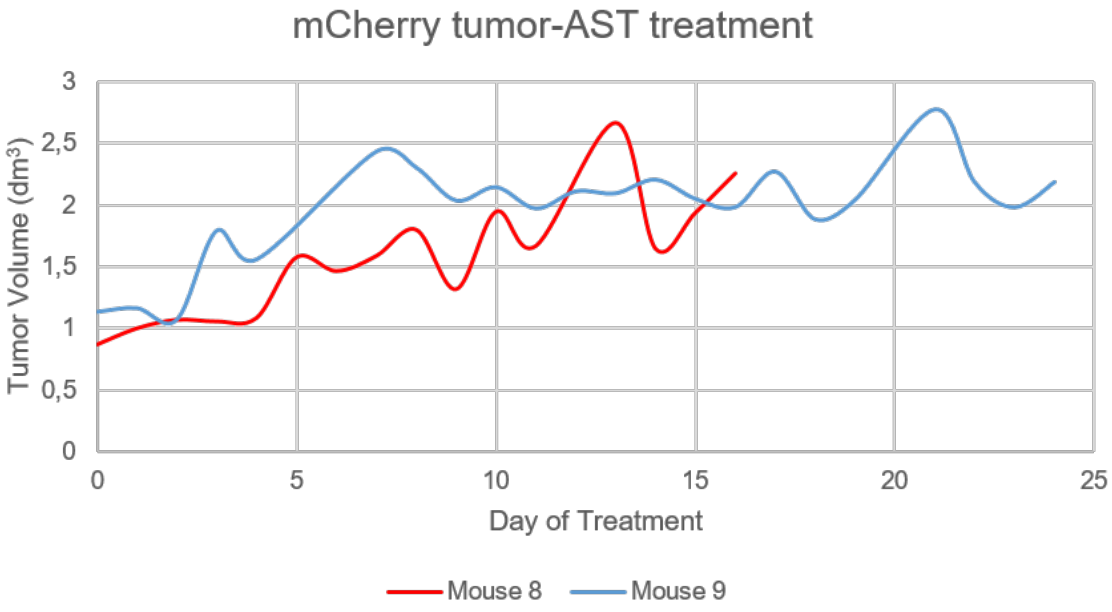


Figure 5.27: mCherry-tumor volumes (dm^3) after AST-C treatment, of Mouse #8 and #9.

The other observations about Mouse #1 were done by a veterinarian. Because the tumors have been reduced very fast, we asked if there may occur other major problems on mouse due to peptide treatment. A check-up list was prepared for all mice (Table 5.6). The only abnormality of Mouse #1 and #2 detected during necropsy was the swollen mesenchymal lymph nodes, which were normal for cancer patients. The other organs did not show apparent signs for other problems.

Table 5.6: Macro pathological conditions of Mouse #1 and #2 before necropsy.

	Mouse 1	**Mouse 2**
Date of check-up	**13.06.2016**	**28.06.2016**
External		
Movements	normal	normal
Response to stimuli	yes	yes
Self cleaning	yes	less
Food-water consumption	normal	normal
Gaita	normal	normal
Opened eyes	open	normal
Nasal discharge	none	none
Ear stream	none	none
Mouth stream	none	none
Color of mucous membrane	pink	pink
Lesion on the skin	none	none
Body temperature	normal	normal
Dehydration	nd	**10%**
Internal		
Gall bladder	normal	normal
Liver	normal	normal contour, normal color
Stomach	normal	normal
Spleen	normal	**grown**
Lymphs	**swollen mesenchymal lymph node**	**swollen**
Kidney	normal	normal, normal borders of medullar cortex
Others	-	normal heart and lungs

5.5.4. Stable Cell Line Generation with a Lentiviral System

AlstR protein and mRNA expression of stable cell lines could not be verified via IRES-GFP construct. Therefore, the generation of new cell lines with a new construct was planned. In order to obtain more reliable and long-lasting expression in Huh7 cells CamAlstR-C gene was cloned into pLENTI-III-HA vector. For control groups, the mCherry sequence was cloned into pLENTI-III-HA vector. Expression of the HA-fused AlstR was confirmed in Western Blot after 24 hr of transient transfection of Huh7 cells (Figure 5.28). *In silico* predicted size of AlstR-C was around 50 kD. However, the obtained mass was between 60 kD and 85 kD. This was the first evidence of the mass of CamAlstR-C protein. Then these constructs were used for virus production in HEK293FT cells, in combination with the other packaging vectors. On the third day after transduction, Huh7 cells were visualized for GFP expression in the GFP control group (Figure 5.29). Simultaneously with the cloning procedures, kill curve of Puromycin on Huh7 cells was constructed and the minimum dose was found as 2 μg/mL. This dose was used for the selection of transduced cells. Another control cell line, such as empty-plasmid-transduction group, was also generated with empty pLENTI-III-HA vector.

5.5.5. Effect of CamAlstRC-ASTC System on Cell Cycle of Huh7 Cells

In the previous studies of cell cycle analysis on previous stable cell lines, AlstR-expressing cells were forming clumps and making FACS analysis more difficult. Therefore, we have performed the same analyses on transiently transfected cells. With the help of new constructs in the following progress duration FACS analysis was replicated (Figure 5.30). The result was the same with the one obtained in previous studies. AST-C treatment did not affect the cell cycle behavior of the empty plasmid-transfected cells. When AlstR was expressed, Huh7 cells lost their G2 phase, leading to an accumulation in G1 and S phases. And when AlstR-expressing cells were treated with AST-C, their cell cycle behavior was restored as in cases of empty plasmid transfection groups (Figure 5.30). In order to understand if the loss of G2 phase was not an outlier due to experimental conditions, optimization experiments were performed on non-transfected

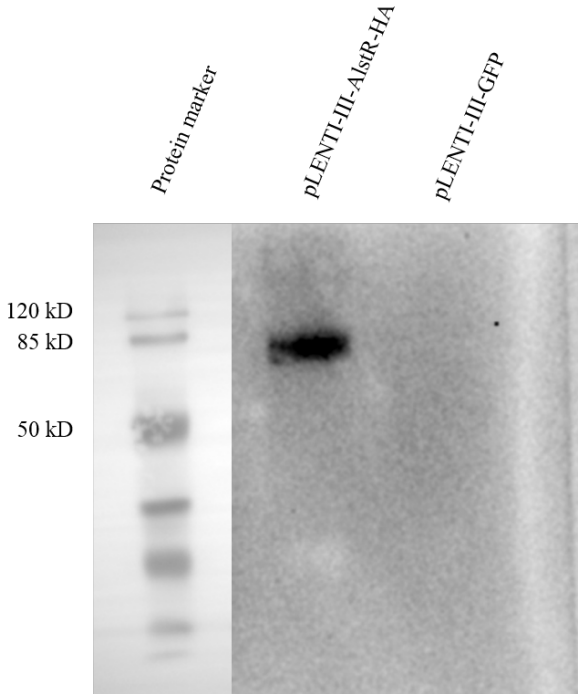


Figure 5.28: Protein expression of HA-fused CamAlstR-C in transiently transfected cells. The WB analysis was performed with anti-HA antibody.

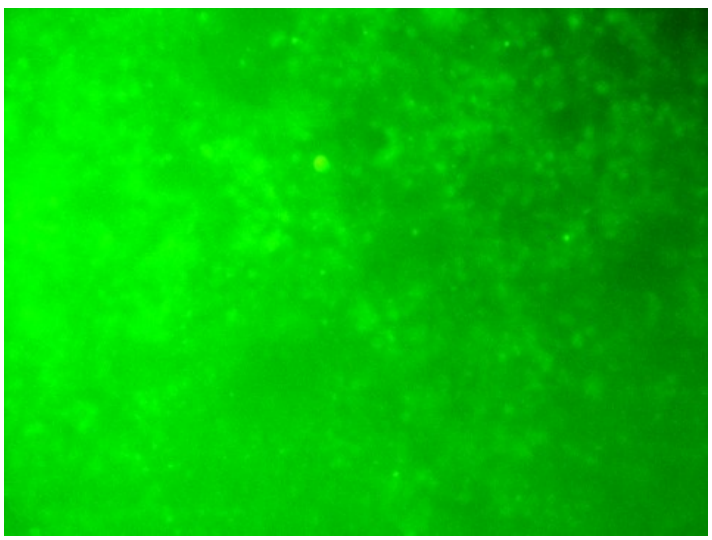


Figure 5.29: Transduction efficiency on third day of transduction of Huh7 cells. The GFP signal was observed for GFP-control cells.

Huh7 cell lines (data not shown). Some parameters such as the type and amount of fixation chemical (EtOH vs MetOH), vortexing during fixation, amount of PI used and the beginning confluency of cells were optimized. Non-transfected Huh7 cells showed a G1 phase of around 40%, S phase of around 40% and G2 phase of around 20% which were also similar with the results of empty-plasmid transfected cells. Additionally, AST-C treatment was performed on non-transfected cells and this treatment did not affect the behavior of the cells (data not shown).

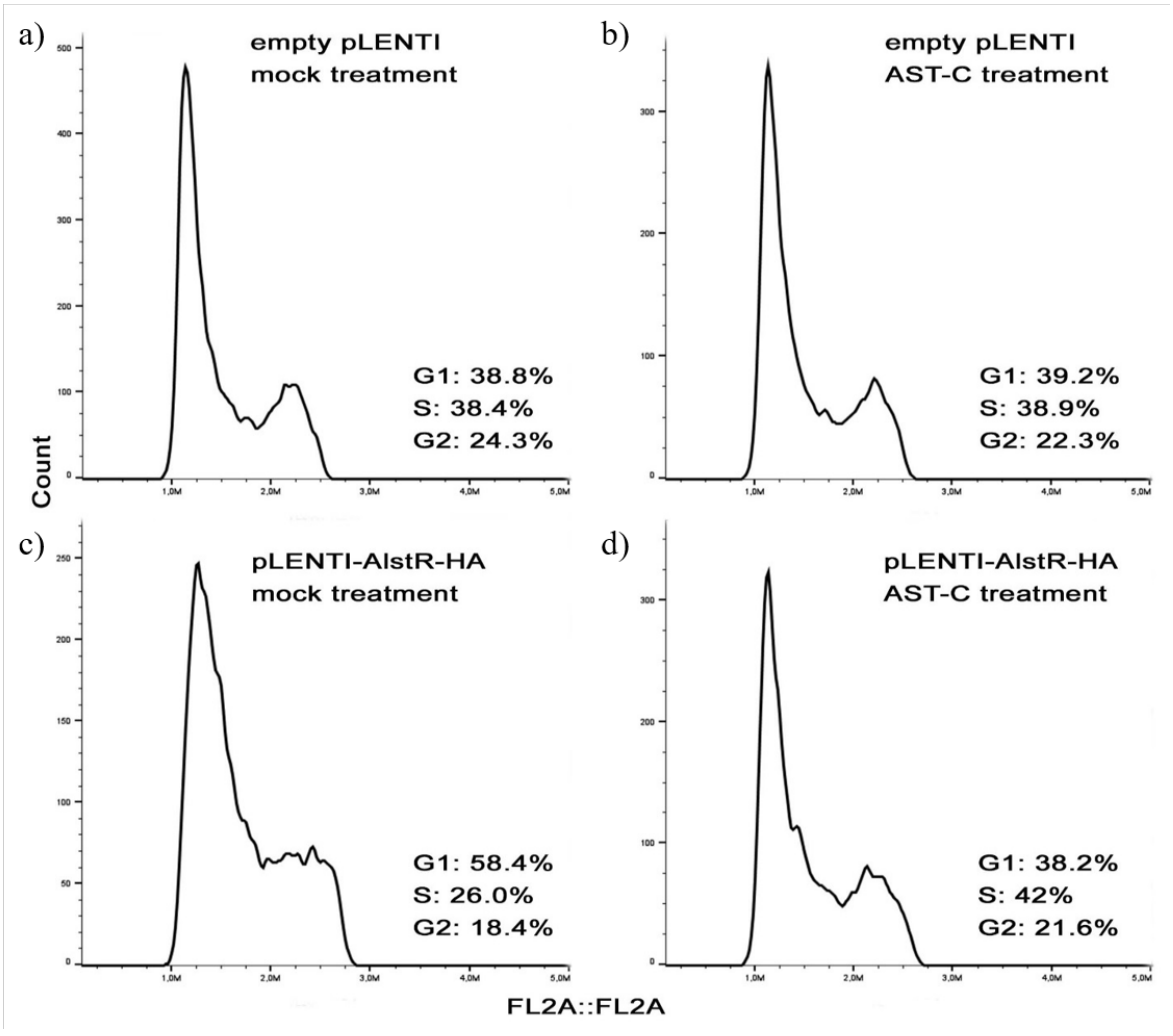


Figure 5.30: PI staining and cell cycle analysis performed on transiently transfected Huh7 cells. Mock treatment group was treated with 0.1% BSA in PBS and AST-C treatment was performed with 10 μM ligand in 0.1% BSA in PBS. The phase quantifications were calculated according to Dean-Jet-Fox model.

5.5.6. Effect of CamAlstRC-ASTC System on Viability of Huh7 Cells

In this part of the study, XTT assays were replicated for 36, 48 and 72 hr treatments on the previous stable cells with the peptide of *D. melanogaster* (Figure 5.31). The response of both AlstR and GFP-expressing cells were similar. Until 5 μM ligand, viability of cells increased slightly. After 5 μM some changes occurred and these changes differed according to the duration of ligand treatment. For instance, at 48 and 72 hr of treatment GFP-expressing cells make a sudden increase in viability at 10 μM treatment. Then the cell viability began decreasing (<100%) after 50 μM treatment. In the case of AlstR-expressing cells, the viability of cells decreased (<100%) slightly after 10 μM treatment. The LC50 values obtained in 36 hr of treatments (497 μM for GFP-stables and 437 μM for AlstR-stables) were much bigger than the values of 48 hr (197 μM for GFP stables and 150 μM for AlstR-stables) and 72 hr treatments (178 μM for GFP-stables and 121 μM for AlstR-stables). However, these LC50 values were very high and the values of GFP-stables were very close to that of AlstR-stables.

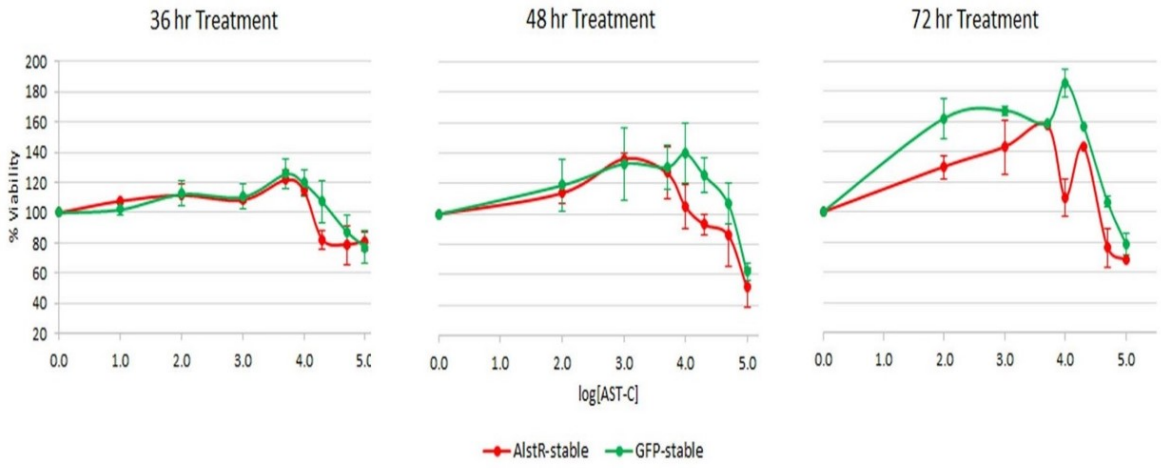


Figure 5.31: XTT viability assays related to AST-C treatment of stable AlstR and GFP-expressing cells, after 36, 48 and 72 hr of AST-C treatment. X-axis contributes to the logarithm of nM AST-C concentrations. Errors (s.e.m.) were calculated from triplicate of 36 hr, triplicate of 48 hr and duplicate of 72 hr treatments.

5.5.7. Effect of AST-C Peptide on Viability of Various Cancer Cells

In order to understand if AST-C can affect Huh7 cells in a receptor-independent way, XTT assays were performed on WT Huh7 cells. Additionally, different cancer cell lines were analyzed simultaneously. Embryonic kidney cells were used as a control to these cancer cell lines. The cells that were used included two types of liver cancer (Huh7 and HepG2), one type of breast cancer (MCF7), one type of ovary cancer (HeLa) and one type of melanoma cancer (MeWo) cell line. The types of that are sensitive to somatostatin analogs, such as pancreatic and lung cancer cell lines, could not be added due to their unavailability. The peptide that was previously used in xenograft, XTT and FACS experiments (DroAST-C) was used again in these XTT experiments together with the newly synthesized CamAST-C peptide. None of the cells responded to both of the peptides in 48 hr (Figure 5.32) or 72 hr of treatments (Figure 5.33).

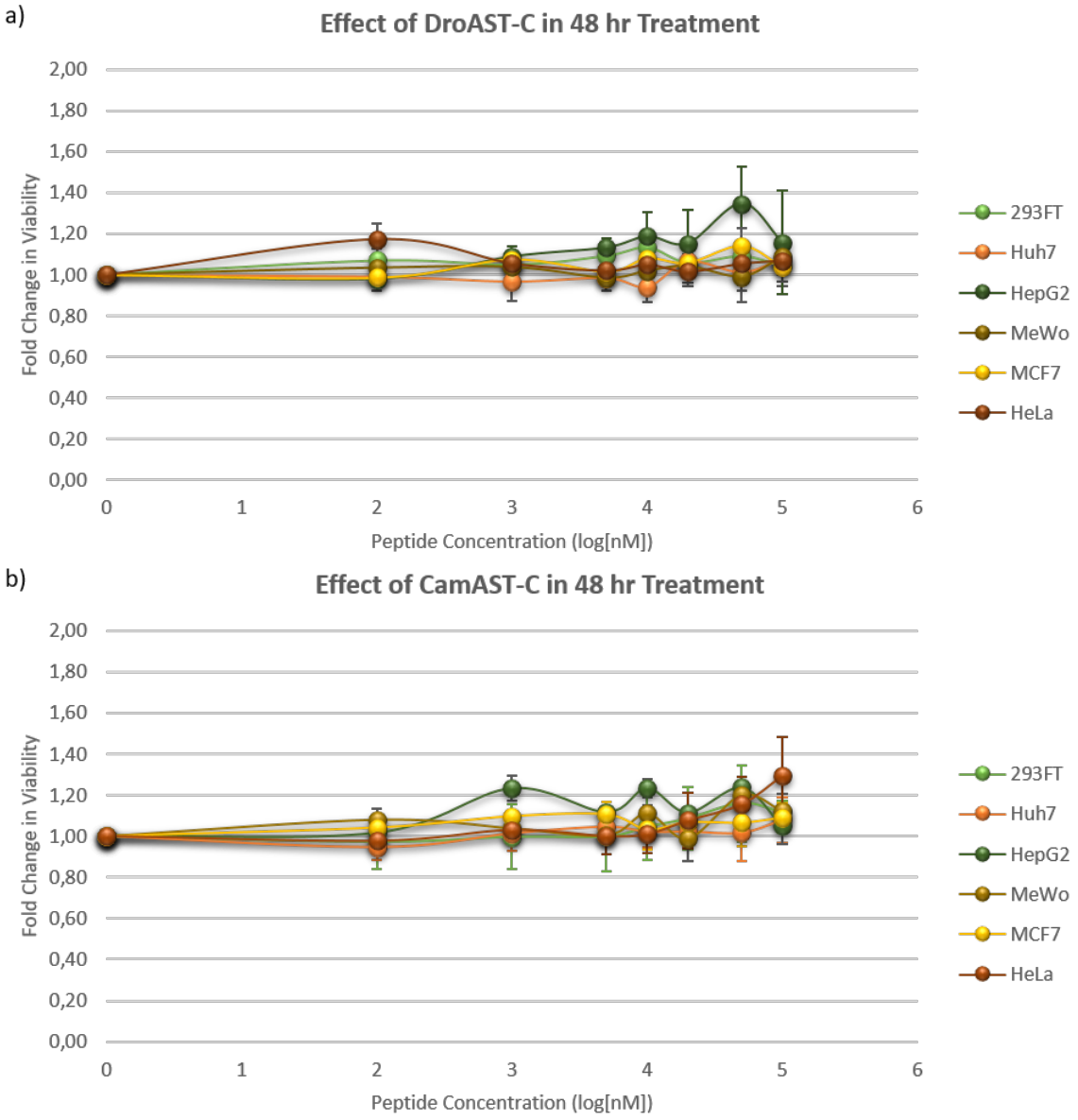


Figure 5.32: XTT cell viability assays for the response of different cancer cell lines after 48 hr treatment of DroAST-C (a) and CamAST-C (b). The absorbances were normalized to mock-treatment group (0 nM). Errors were calculated as Standard Error of The Mean (n=3).

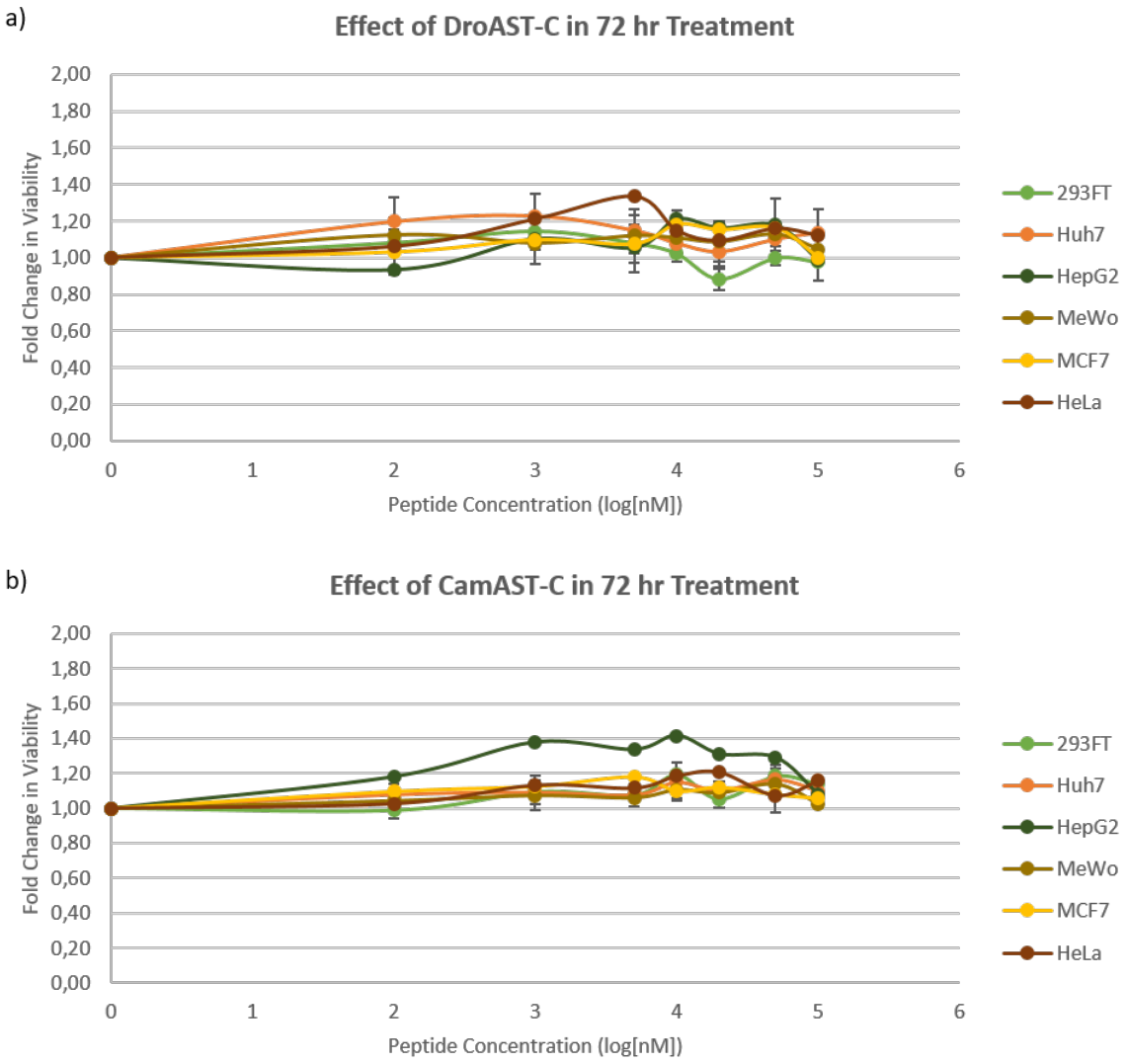


Figure 5.33: XTT cell viability assays for the response of different cancer cell lines after 72 hr treatment of DroAST-C (a) and CamAST-C (b). The absorbances were normalized to mock-treatment group (0 nM). Errors were calculated as Standard Error of The Mean (n=3).

6. CONCLUSION AND DISCUSSION

6.1. IXTPP Motif on ECL3 is Important for the Ligand Binding of CamAlstR-C

The important elements of binding pocket of a novel GPCR, CamAlstR-C, were verified in this study. The results showed that IXTPP motif which was conserved within ECL3 regions of insect AlstRs was essential in ligand binding of CamAlstR-C. In addition, variable N-terminal loops of these GPCRs seemed another important structural element in this binding. Therefore, a reliable information for future agonist/antagonist studies could be obtained.

Another result was obtained from AFM experiments with AST-C peptide. In higher loading rates (>106 pN/sec) AST-C unbinding forces showed a sharp increase, changing the equation of the curve. This may be an indication of two-step energy barrier in this unbinding event as reviewed by Evans, Williams and Lee [38]; [39]. The complexity of this unbinding may stem from i) breakage of the disulfide bond of AST-C, ii) conformational change on CamAlstR-C or iii) flexibility of binding pocket of CamAlstR-C.

In addition to these results, the method used in this part proposed a valuable tool for binding pocket studies. Here we have combined computational tools with single-molecule force spectroscopy in which we could directly measure binding strengths in physiological conditions. GPCRs are not widely used in structural studies such as X-ray and NMR and they behave different when they are not engaged to a lipid membrane or even when they are not in physiological conditions. Therefore, our combined computational-experimental approach can serve as an easy and reliable method in drug design studies of GPCRs.

In our laboratory, some functional assays were conducted (by Ali İşbilir, Molecular Biology and Genetics, Boğaziçi University) with the same mutant CamAlstR-C forms.

In these experiments both FRET and TGF-α Shedding Assay measurements were performed. As a result, AAAAA mutant form of the receptor showed lower EC50 values than the wild type receptor and other mutant receptor forms [8]. In conclusion of this part, we supported that the IXTPP motif in the binding pocket of novel AlstR-C of *C. morosus* was crucial for binding with the ligand and for activation with the ligand.

6.2. The Ligand of CamAlstR-C was Obtained From the Neuropeptidome, as well as the Other Neuropeptides Expressed in Adult Animal.

The neuropeptides and their cognate GPCRs play crucial roles in many aspects of insects, such as development, social behavior, feeding behavior, sleep cycles, egg laying or other physiological events. Evolutionarily they became the main signal transduction elements since one of the earliest organisms *Hydra*. These molecules represent a very big family and valuable targets for many studies. Before this study, there was no information about the original ligand of CamAlstR-C. Therefore, AST-C peptides of other organisms, such as *Drosophila*, were utilized in the assays. It is anticipated that the AST-C of *C. morosus* might change the outcome of the quantitative assays. As a result, the subsequent part of the study focused on determination of the original ligand of CamAlstR-C and other three types of AlstRs in *C. morosus*.

In order to achieve this goal, RNA sequencing was performed and this yielded a large amount of information about all of the GPCRs and neuropeptides expressed in adult *C. morosus* tissues. Twenty nine putative neuropeptide precursor sequences were obtained from RNA assembly, which gave rise to twenty three types of neuropeptides. *C. morosus* belongs to the order of Phasmatodea and three closest species which have genome assembly data are *Blatella germanica* (German cockroach), *Locusta migratoria* (Migratory locust) and *Zootermopsis nevadensis* (termite). Among these species, migratory locust is known to express 44 neuropeptides [40] and the termite 59 neuropeptide precursors [41]. On the other hand, 80 genes were identified in human expressing at least 150 different mature neuropeptides. Also, in *Drosophila* genome, 42 genes were identified to code for neuropeptides [42]. Our number is small compared

to the closest relatives. Because of the fact that we could predict only the expressed neuropeptides in adult tissues, this number should be less than the number which is predicted from a genome.

The most important finding of this study is the mature sequence of original peptide of CamAlstR-C. And for the following studies, we used the commercially synthesized peptide in cell culture and *in vivo* experiments. The presence of other transcripts indicated that AST-A and AST-CC peptides were also expressed in this organism.

In the following part of the neuropeptidome search, we aimed to perform proteomics analysis via nano-LCMS technique. However on April 2018 the neuropeptidome of *C. morosus* has been published. Liessem and his colleagues in University of Cologne combined MALDI-TOF MS and nano-LCMS to detect even the low abundant peptides in different neuronal segments (separate ganglia) [3]. They could detect 60 propeptide sequences together with 5 novel mature neuropeptides. Our data has showed some consistencies and differences with their data. For instance, they showed that there were 12 allatostatin A (AST-A) peptides, but in our prediction one AST-A propeptide could lead to 13 mature peptides. When we checked the presence of the neuropeptides which were undetected in our analysis but present in Liessem's work, 14 of these peptides could be detected in our analysis also. These peptides are agotoxin, calcitonin b, IDL containing peptide, ITG-like peptide, myoinhibitory peptide a, NVP-like a1 and 2, PKL1 and 2, proctolin, RFLamide, trissin, tryptopyrokinin and their novel peptide hansolin. However, 25 neuropeptides in their results could not be detected in our transcriptome. The reason for us to be unable to find the 14 neuropeptides before might be the queries that we have chosen. If the queries included in the initial search were not similar we could have not found any hits.

6.3. GPCRome of *C. morosus* Gave Clues About the Functions of Specific GPCRs.

Presence of the other allatostatin receptors was the second question that was aimed to be extracted from RNAseq data. With the help of Blastx tool, we could find

out the partial mRNA sequences for AlstR-A and AlstR-C. The literature on AlstR-B sequence was inadequate in databases and it was used as synonymous with the myoinhibitory peptide receptor (MIPR). So, we used the MIPR sequences to find a putative AlstR-B transcript in our transcriptome. We could detect partial similar transcripts but the similarity was not significant. So we could not conclude that AlstR-B or MIPR was also expressed in adult *C. morosus*. Expression of CamAlstR-C and CamAlstR-A revealed slightly uniform distribution within different tissues. The results of Liessem *et al.* showed the presence of AST-A, allatotropin, myoinhibitory peptide, small neuropeptide F and other peptides in the frontal ganglion which regulates the motility of foregut [3]. Our data showed that CamAlstR-A was abundant in the head and the foregut which may be the target of AST-A peptide secreted from frontal ganglion. Secondly, the literature on AST-A shows its presence and inhibitory role on contraction of the hindgut [43]. The expression levels in our data was lower in hindgut than in foregut or gastric cecea, but it was still expressed. CamAlstR-C is also uniformly expressed but having the highest levels in brain, CC and CA samples. This result is consistent with the data of other insects such as mosquito [44], with the expression of AST-C in frontal ganglion [3] and with its function. AlstR-A and AlstR-C receptors are both named because they inhibit Juvenile Hormone synthesis and secretion from CA. However, this function depends on the species. For instance in mosquitos AST-C inhibits the JH synthesis but AST-A does not [45]. On the other hand, AST-A is the inhibitor of JH synthesis in *Diploptera punctata* and *Periplaneta americana* [46]. With the help of our data we can predict that both types of allatostatins can have important roles in different functions, but it is possible that C-type can be the inhibitor of JH more strongly than A-type, due to its expression profile in the brain and neuroendocrine glands.

Expression profiles of most of the other GPCRs revealed expected results. For example, AKHR was expected to be expressed in fat body and in the head [47]. Its ligand was found in the proteomic analysis of CC [3]. And in our results, it was highly expressed in fat body as well as ganglia.

DHR was expected to be highly expressed in Malpighian tubules due to its functions in water homeostasis [48]. Our results did not show a significant difference but the highest expression was in Malpighian tubules in accordance with the literature. A similar receptor was CamCalR which has functions in calcium homeostasis. Its expression depends on the species [49] and in *C. morosus* it was significantly expressed in fat body.

CCKR is also called as substance P receptor and the homolog of human tachykinin receptor which is responsible for the stress and pain responses in human [50]. In insects, it has similar roles such as the aggressive behavior of *Drosophila* [51]. It was mainly expressed in the central nervous system and the gut [52]. Our results supported its presence in the parts of the gut, but its expression in the head was lower than expected. Inotocin peptide was absent in the proteomic data [3]. However, we could detect significant expression of its receptor in gastric cecea (including the anterior midgut). It's the insect homolog of vasopressin/oxytocin family peptides and has roles in reproductive behavior of the animals [53].

Frizzled 10 expression level was very low when compared to ovary levels probably because of its high expression in the mature eggs inside the ovary. Still in the ovary, its expression level was much lower than the other GPCRs. This is probably due to the developmental functions of this receptor.

Octopamine and tyramine receptors were expected to be abundant primarily in CNS, and then in intestine, Malpighian tubules and also other organs [54], but the results were not compatible with the literature of other insects. In our results, CamTyr2R was highly expressed in aorta, but CamOctR didn't show a tissue-specific expression profile.

Gustatory receptor for sugar taste 43a is a recently identified taste receptor which is mostly expressed in brain as well as the gastrointestinal tract [55]. Our data couldn't support its presence in the brain, but it was expressed in gastric cecea.

Adhesion GPCRs are not well studied in insects. Especially adhesion GPCR G2 was found to be functional in reproductive system [56]. However, in our study CamAdgrG2 showed that it was expressed in ganglia and the parts of the gut more than other tissues, but with insignificant difference. And at least two types of adhesion GPCRs (G2 and A3) were expressed in this organism. This result may serve as a start for future studies on adhesion GPCRs of insects.

NPYRs are activated by sNPF peptides and responsible for various functions such as appetite [57] or circadian rhythm, [58]. And in proteomics of *C. morosus*, sNPF was present in CC. Our results exhibited lower levels of expression in any of the tissues than ovary, but the second and third tissues expressing NPYR were brain, CC, CA and ganglia.

The most important result came from an orphan receptor. Within the transcriptome data, a partial mRNA sequence was detected to be similar to an uncharacterized GPCR. In the expression analysis, it showed specific expression in brain, CC, CA and ganglia samples. This result can facilitate further studies on deorphanization of this GPCR.

In conclusion, we could reveal the types of GPCRs that are expressed in adult *C. morosus* body and then their expression profiles in different tissues. These results can help further studies on characterization of these GPCRs in this organism.

6.4. AlstRC-ASTC System has no Effect on Proliferation of Cancer Cells.

AlstR-C is the homolog of human somatostatin receptor (SSTR). And the SSTRs have anti-proliferative roles and are currently studied as targets in cancer research. When compared to SSTR mechanism, it was hypothesized that activation of AlstR could also lead to a cell cycle arrest or apoptosis. ERK phosphorylation response was also the effect of SSTR-SST activity on breast cancer cells [59]. And its nuclear localization could lead to an apoptotic response in these cells. This hypothesis lead us to study if AlstR-C may exert the same effects on mammalian cells. The first question

in this study was to see if a constitutively active AlstR-C could be utilized against cancer growth. However, in order to answer this question we needed to test if the ligand-activated AlstR-C could have anti-proliferative abilities or not.

The earliest results on *Drosophila* nerve cell lines showed that AST-C treatment lead to a growth arrest similar to the effects in absence of growth stimulating insulin hormone. In addition, when we ectopically express CamAlstR-C in cancer cell lines and treat them with AST-C peptide, the G2 phase arrest was observed as in *Drosophila* cell lines. Viability of CamAlstR-C-expressing and GFP-expressing Huh7 cells was also decreased with increasing AST-C treatments, but both cells gave similar responses. Then we wanted to try this receptor-ligand system on mouse tumor xenografts. This hypothesis was again based on their homology to SSTR-SST system of mammalian cells. SST analogs were used against NETs, but they were inefficient, having side-effects and should be used in combination with other adjuvants. Because AlstR is not expressed in mammalian cells, an AlstR activation in targeted cells could result in an effect which would not harm the health of non-targeted cells. And if we could increase the activity of AlstR, it could give more efficient results than that of SST analogs. In order to see its long-term effects, we planned *in vivo* experiments.

The first mouse xenografts gave promising results. AST treatment reduced the size of both tumors (GFP and AlstR-expressing tumors) when compared to Saline treatment. This reduction began after seventh-eighth day of treatment. However, only the AlstR-expressing tumor was supposed to be reduced with AST treatment because AST peptide should not be interacting with any other surface receptor of Huh7 cells. The reduction in GFP-expressing tumor was unexpected and other questions occurred: i) is the effect of AST specific for AlstR-tumors? and ii) does AST negatively affect the own cells of the animals? For the second question, macro clinical investigations were noted and the only differences on these mice were swallowing of the lymph nodes (which is normal for cancer patients) and mild dehydration. In order to understand whether this peptide can stimulate a response independent of its receptor, other tumor xenografts were performed with mCherry-expressing and wild type Huh7 cell lines. The other observation on the tumors was the absence of veins inside and outside of

the tumor capsule. This result lead to a third question about an anti-angiogenic effect because one of the effects of SSTRs was shown as anti-angiogenesis [19]. The tumors might have shrunken due to nutrient deprivation. Therefore, we proposed a two-step response against AST-C treatment: 1) cell cycle arrest which was observed in cultured cells and 2) cell death upon continuous treatment, which may be an indirect effect of inhibition of angiogenesis. However, at first we needed to confirm that AST-C behaved like an anti-proliferative agent on cancer cells.

The results obtained with mCherry-expressing and wild type Huh7 xenografts were not reliable. The number of tumors that grew on mice was inadequate, so saline treatments could not be replicated. Tumor volumes and growth rates were not consistent. For instance, AlstR-expressing tumor of Mouse #2 was smaller at the beginning of AST treatment but grew faster than that of Mouse #5. In contrast, GFP tumor of Mouse #5 was smaller than that of Mouse #2 but it did not grow while that of Mouse #2 grew very fast. AlstR-expressing tumor of Mouse #4 grew more slowly than its GFP-expressing tumor, while GFP-expressing tumor of Mouse #5 was stabilized more than its AlstR-expressing tumor. We could not conclude any suggestion about the efficiency of AST on these tumors. Nevertheless, mCherry-expressing tumor of Mouse #9 exhibited no growth after 7-8 days of treatment and one piece of the tumor was almost disappeared on the 23rd day of treatment. This observation was similar to that observed in GFP-expressing tumor of Mouse #1. These responses coming from AlstR-free tumors strengthened the questions about specific activity of AST-C against AlstR-C. However, the last mCherry-expressing tumor did not give the same result. Additionally, the following xenografts showed that WT-Huh7 tumors did not shrink at all. The numbers of each type of tumor were at least 3, except saline treatment groups of stable cell line tumors (AlstR, GFP and mCherry). Therefore, we can say that the promising results of the first tumors could not be reproduced.

We replicated the cell cycle analysis via different constructs. The results showed that AST-C treatment did not affect the cell cycle progression of empty plasmid transfected cells. If the cells express CamAlstR-C without ligand treatment, they lost their G2 phase, most probably due to an arrest in G1 and S phases. Unexpectedly, ligand

treatment of these cells made them to recover their G2 phase. This response was unexpected because in theory a receptor can exert its basal activity when overexpressed in cells, but this activity should increase with ligand treatment. However, in our experimental set-up the basal activity of the receptor lead to loss of G2 and ligand activation reverted this effect. This behavior was similar to the results obtained in the previous experiments.

The responses of AlstR, GFP and mCherry-expressing cells were very consistent both in xenograft models and XTT assays. They responded to AST-C treatment, but the WT cells did not respond. Even the other types of cancer cell lines did not lose viability in XTT assays. All these results revealed other questions. What might be happening in the transfected cells so that they respond to AST-C peptide even in the absence of its receptor? Or is over-expression of GFP or mCherry causing a stress in cells that results in a sensitivity to AST peptide? At the end, we conclude that AST-C treatment exerted no effect on the proliferative abilities of various cancer cell lines even if the cells express its cognate GPCR.

Although a possible anti-proliferative effect has not been confirmed in this part of the study, this study provided us with a lot of information about AlstR-C and AST-C. The results obtained so far are very important both in terms of combining *in silico* and nano-scale studies as well as in the literature on insect neuropeptides and GPCRs. And in the future, it has made it possible to gather preliminary information about a molecule that is likely to become more popular.

REFERENCES

1. Duan Sahbaz, B. and N. Birgul Iyison, "Neuropeptides as Ligands for GPCRs", C. Saravanan (Editor), *Ligands*, chap. 4, pp. 77–101, IntechOpen, 2018.

2. Bläsing, B. and H. Cruse, "Mechanisms of stick insect locomotion in a gap-crossing paradigm", *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, Vol. 190, No. 3, pp. 173–183, 2004.

3. Liessem, S., L. Ragionieri, S. Neupert, A. Buschges and R. Predel, "Transcriptomic and Neuropeptidomic Analysis of the Stick Insect, Carausius morosus.", *Journal of proteome research*, p. acs.jproteome.8b00155, 2018.

4. Birgül, N., C. Weise, H. J. Kreienkamp and D. Richter, "Reverse physiology in drosophila: identification of a novel allatostatin-like neuropeptide and its cognate receptor structurally related to the mammalian somatostatin/galanin/opioid receptor family.", *The EMBO journal*, Vol. 18, No. 21, pp. 5892–900, nov 1999.

5. Audsley, N., H. J. Matthews, N. R. Price and R. J. Weaver, "Allatoregulatory peptides in Lepidoptera, structures, distribution and functions.", *Journal of insect physiology*, Vol. 54, No. 6, pp. 969–80, jun 2008.

6. Martín, D., M. D. Piulachs and X. Bellés, "Inhibition of vitellogenin production by allatostatin in the German cockroach.", *Molecular and Cellular Endocrinology*, Vol. 121, No. 2, pp. 191–196, 1996.

7. Duan Sahbaz, B., O. U. Sezerman, H. Torun, &. Necla and B. Iyison, "Ligand binding pocket of a novel Allatostatin receptor type C of stick insect, Carausius morosus", *Scientific Reports*, Vol. 7, 2017.

8. İsbilir, A., *Pharmacological characterization of PISCF-Allatostatin Receptor in Carausius morosus*, Master's Thesis, Bogazici University, 2016.

9. Murck, H., I. A. Antonijevic, R. M. Frieboes, P. Maier, T. Schier and A. Steiger, "Galanin has REM-sleep deprivation-like effects on the sleep EEG in healthy young men", *Journal of Psychiatric Research*, Vol. 33, No. 3, pp. 225–232, 1999.

10. Tan, E. M., Y. Yamaguchi, G. D. Horwitz, S. Gosgnach, E. S. Lein, M. Goulding, T. D. Albright and E. M. Callaway, "Selective and Quickly Reversible Inactivation of Mammalian Neurons In Vivo Using the Drosophila Allatostatin Receptor", *Neuron*, Vol. 51, pp. 157–170, 2006.

11. Nielsen, K. J., E. M. Callaway and R. J. Krauzlis, "Viral vector-based reversible neuronal inactivation and behavioral manipulation in the macaque monkey", *Frontiers in Systems Neuroscience*, Vol. 6, No. June, pp. 1–12, 2012.

12. Yamada, K., E. Wada and K. Wada, "Bombesin-like peptides: studies on food intake and social behaviour with receptor knock-out mice", *Annals of Medicine*, Vol. 32, No. 8, pp. 519–529, 2000.

13. Simo, L., J. Koci and Y. Park, "Receptors for the neuropeptides, myoinhibitory peptide and SIFamide, in control of the salivary glands of the blacklegged tick Ixodes scapularis", *Insect Biochemistry and Molecular Biology*, Vol. 43, No. 4, pp. 376–387, 2013.

14. Stepanyan, Z., A. Kocharyan, M. Behrens, C. Koebnick, M. Pyrski and W. Meyerhof, "Somatostatin, a negative-regulator of central leptin action in the rat hypothalamus", *Journal of Neurochemistry*, Vol. 100, No. 2, pp. 468–478, 2007.

15. Olias, G., C. Viollet, H. Kusserow, J. Epelbaum and W. Meyerhof, "Regulation and function of somatostatin receptors", *Journal of Neurochemistry*, Vol. 89, No. 5, pp. 1057–1091, 2004.

16. Brazeau, P., W. Vale, R. Burgus, N. Ling, M. Butcher, J. Rivier and R. Guillemin, "Hypothalamic polypeptide that inhibits the secretion of immunoreactive pituitary growth hormone.", *Science*, Vol. 179, No. 68, pp. 77–79, 1973.

17. Ferjoux, G., C. Bousquet, P. Cordelier, N. Benali, F. Lopez, P. Rochaix, L. Buscail and C. Susini, "Signal transduction of somatostatin receptors negatively controlling cell proliferation.", *Journal of physiology, Paris*, Vol. 94, No. 3-4, pp. 205–10, 2000.

18. Florio, T., S. Thellung, S. Arena, a. Corsaro, a. Bajetto, G. Schettini and P. J. Stork, "Somatostatin receptor 1 (SSTR1)-mediated inhibition of cell proliferation correlates with the activation of the MAP kinase cascade: role of the phosphotyrosine phosphatase SHP-2.", *Journal of physiology, Paris*, Vol. 94, No. 3-4, pp. 239–50, 2000.

19. Kumar, U. and M. Grant, "Somatostatin and somatostatin receptors", *Results and Problems in Cell Differentiation*, Vol. 50, pp. 137–84, 2010.

20. He, Y., X.-M. Yuan, P. Lei, S. Wu, W. Xing, X.-L. Lan, H.-F. Zhu, T. Huang, G.-B. Wang, R. An, Y.-X. Zhang and G.-X. Shen, "The antiproliferative effects of somatostatin receptor subtype 2 in breast cancer cells.", *Acta pharmacologica Sinica*, Vol. 30, No. 7, pp. 1053–9, July 2009.

21. Ruscica, M., M. Arvigo, F. Gatto, E. Dozio, D. Feltrin, M. D. Culler, F. Minuto, M. Motta, D. Ferone and P. Magni, "Regulation of prostate cancer cell proliferation by somatostatin receptor activation.", *Molecular and cellular endocrinology*, Vol. 315, No. 1-2, pp. 254–62, February 2010.

22. Bousquet, C., C. Duluc, S. Moatassim-Billah, Y. Bakri, M.-B. Delisle, H. Schmid, M. Mathonnet and S. Pyronnet, "Antimetastatic potential of somatostatin analog SOM230 (pasireotide) in pancreatic cancer: Indirect pharmacological targeting of cancer-associated fibroblasts", *Pancreatology*, Vol. 15, No. 3, pp. S32–S33, 2015.

23. Toumpanakis, C. and M. E. Caplin, "Update on the Role of Somatostatin Analogs for the Treatment of Patients With Gastroenteropancreatic Neuroendocrine Tumors", *Seminars in Oncology*, Vol. 40, No. 1, pp. 56–68, 2013.

24. Wolin, E. M., B. Jarzab, B. Eriksson, T. Walter, C. Toumpanakis, M. A. Morse,

P. Tomassetti, M. M. Weber, D. R. Fogelman, J. Ramage, D. Poon, B. Gadbaw, J. Li, J. L. Pasieka, A. Mahamat, F. Swahn, J. Newell-Price, W. Mansoor and K. Öberg, "Phase III study of pasireotide long-acting release in patients with metastatic neuroendocrine tumors and carcinoid symptoms refractory to available somatostatin analogues", *Drug Design, Development and Therapy*, Vol. 9, pp. 5075–5086, 2015.

25. Singh, B., J. A. Smith, D. M. Axelrod, P. Ameri, H. Levitt, A. Danoff, M. Lesser, C. de Angelis, I. Illa-Bochaca, S. Lubitz, D. Huberman, F. Darvishian and D. L. Kleinberg, "Insulin-like growth factor-I inhibition with pasireotide decreases cell proliferation and increases apoptosis in pre-malignant lesions of the breast: a phase 1 proof of principle trial.", *Breast cancer research : BCR*, Vol. 16, No. 6, p. 463, 2014.

26. Dasgupta, P. and R. Mukherjee, "Lipophilization of somatostatin analog RC-160 with long chain fatty acid improves its antiproliferative and antiangiogenic activity in vitro.", *British journal of pharmacology*, Vol. 129, pp. 101–109, 2000.

27. Ferrante, E., C. Pellegrini, S. Bondioni, E. Peverelli, M. Locatelli, P. Gelmini, P. Luciani, A. Peri, G. Mantovani, S. Bosari, P. Beck-Peccoz, A. Spada and A. Lania, "Octreotide promotes apoptosis in human somatotroph tumor cells by activating somatostatin receptor type 2", *Endocrine-Related Cancer*, Vol. 13, No. 3, pp. 955–962, 2006.

28. Down, R. E., H. J. Matthews and N. Audsley, "Effects of Manduca sexta allatostatin and an analog on the pea aphid Acyrthosiphon pisum (Hemiptera: Aphididae) and degradation by enzymes from the aphid gut.", *Peptides*, Vol. 31, No. 3, pp. 489–97, March 2010.

29. Auerswald, L., N. Birgül, G. Gäde, H. J. Kreienkamp and D. Richter, "Structural, functional, and evolutionary characterization of novel members of the allatostatin receptor family from insects.", *Biochemical and biophysical research communications*, Vol. 282, No. 4, pp. 904–9, April 2001.

30. Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen and T. L. Madden, "Primer-BLAST: A tool to design target-specific primers for polymerse chain reaction.", *BMC Bioinformatics*, Vol. 18, p. 13:134, 2012.

31. Duan Sahbaz, B., *Identification and Docking Analysis of Allatostatin Receptor in Carausius morosus*, Master's Thesis, Boğaziçi University, 2013.

32. Nielsen, H., "Predicting Secretory Proteins with SignalP", *Methods in Molecular Biology*, Vol. 1611, pp. 59–73, 2017.

33. Duckert, P., S. Brunak and N. Blom, "Prediction of proprotein convertase cleavage sites", *Protein Engineering, Design and Selection*, Vol. 17, No. 1, pp. 107–112, 2004.

34. Ferrè, F. and P. Clote, "DiANNA: A web server for disulfide connectivity prediction", *Nucleic Acids Research*, Vol. 33, No. SUPPL. 2, pp. 230–232, 2005.

35. Krogh, a., B. Larsson, G. von Heijne and E. L. Sonnhammer, "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.", *Journal of molecular biology*, 2001.

36. Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman, "Basic local alignment search tool.", *Journal of molecular biology*, 1990.

37. Isberg, V., S. Mordalski, C. Munk, K. Rataj, K. Harpsøe, A. S. Hauser, B. Vroling, A. J. Bojarski, G. Vriend and D. E. Gloriam, "GPCRdb: An information system for G protein-coupled receptors", *Nucleic Acids Research*, 2016.

38. Evans, E. and P. Williams, "Dynamic Force Spectroscopy", F. Flyvbjerg, F. Jülicher, P. Ormos and F. David (Editors), *Physics of bio-molecules and cells. Physique des biomolécules et des cellules*, pp. 145–204, Springer Berlin Heidelberg, Heidelberg, 2002.

39. Lee, C. K., Y. M. Wang, L. S. Huang and S. Lin, "Atomic force microscopy: Determination of unbinding force, off rate and energy barrier for protein-ligand interaction", *Micron*, Vol. 38, pp. 446–46, 2007.

40. Loof, A. D., L. Schoofs, G. M. Holman, T. Hayes, R. Nachman and V. J. Broeck, "Isolation and Identification of Neuropeptides in Locusta Migratoria", *Chromatography and Isolation of Insect Hormones and Pheromones*, pp. 205–211, Springer Link, 1991.

41. Veenstra, J. A., "The contribution of the genomes of a termite and a locust to our understanding of insect neuropeptides and neurohormones", *Frontiers in Physiology*, Vol. 5, No. Nov, pp. 1–22, 2014.

42. Hewes, R. S. and P. H. Taghert, "Neuropeptides and Neuropeptide Receptors in the Drosophila melanogaster Genome", *Genome Biology*, Vol. 11, pp. 1126–1142, 2001.

43. Secher, T., C. Lenz, G. Cazzamali, G. Sørensen, M. Williamson, G. N. Hansen, P. Svane and C. J. Grimmelikhuijzen, "Molecular Cloning of a Functional Allatostatin Gut/Brain Receptor and an Allatostatin Preprohormone from the Silkworm Bombyx mori", *Journal of Biological Chemistry*, Vol. 276, No. 50, pp. 47052–47060, 2001.

44. Mayoral, J. G., M. Nouzova, A. Brockhoff, M. Goodwin, S. Hernandez-Martinez, D. Richter, W. Meyerhof and F. G. Noriega, "Allatostatin-C receptors in mosquitoes.", *Peptides*, Vol. 31, No. 3, pp. 442–50, March 2010.

45. Li, Y., S. Hernandez-Martinez and F. G. Noriega, "Inhibition of juvenile hormone biosynthesis in mosquitoes: Effect of allatostatic head factors, PISCF- and YXFGL-amide-allatostatins", *Regulatory Peptides*, Vol. 118, No. 3, pp. 175–182, 2004.

46. Lloyd, G. T., a. P. Woodhead and B. Stay, "Release of neurosecretory granules

within the corpus allatum in relation to the regulation of juvenile hormone synthesis in Diploptera punctata.", *Insect biochemistry and molecular biology*, Vol. 30, No. 8-9, pp. 739–46, 2000.

47. Alves-Bezerra, M., I. F. De Paula, J. M. Medina, G. Silva-Oliveira, J. S. Medeiros, G. Gäde and K. C. Gondim, "Adipokinetic hormone receptor gene identification and its role in triacylglycerol metabolism in the blood-sucking insect Rhodnius prolixus", *Insect Biochemistry and Molecular Biology*, Vol. 69, pp. 51–60, 2016.

48. Paluzzi, J.-P., Y. Park, R. J. Nachman and I. Orchard, "Isolation, expression analysis, and functional characterization of the first antidiuretic hormone receptor in insects", *Proceedings of the National Academy of Sciences*, Vol. 107, No. 22, pp. 10290–10295, 2010.

49. Zandawala, M., S. Li, F. Hauser, C. J. P. Grimmelikhuijzen and I. Orchard, "Isolation and functional characterization of calcitonin-like diuretic hormone receptors in Rhodnius prolixus", *PLoS ONE*, 2013.

50. Ebner, K., S. Sartori and N. Singewald, "Tachykinin Receptors as Therapeutic Targets in Stress-Related Disorders", *Current Pharmaceutical Design*, Vol. 15, No. 14, pp. 1647–1674, 2009.

51. Asahina, K., K. Watanabe, B. Duistermars, E. Hoopfer, C. Gonzalez, E. Eyjolfsdottir, P. Perona and D. Andersen, "Tachykinin-expressing neurons control male-specific aggressive arousal in Drosophila", *Cell*, Vol. 156, No. 0, pp. 221–235, 2014.

52. Xu, G., G.-x. Gu, Z.-w. Teng, S.-f. Wu, J. Huang and Q.-s. Song, "Identification and expression profiles of neuropeptides and their G protein-coupled receptors in the rice stem borer Chilo suppressalis", *Nature Publishing Group*, Vol. 6, No. 1, p. 28976, 2016.

53. Chérasse, S. and S. Aron, "Measuring inotocin receptor gene expression in chronological order in ant queens", *Hormones and Behavior*, 2017.

54. El-Kholy, S., F. Stephano, Y. Li, A. Bhandari, C. Fink and T. Roeder, "Expression analysis of octopamine and tyramine receptors in Drosophila", *Cell and Tissue Research*, Vol. 361, No. 3, pp. 669–684, 2015.

55. Mishra, D., T. Miyamoto, Y. H. Rezenom, A. Broussard, A. Yavuz, J. Slone, D. H. Russell and H. Amrein, "The molecular basis of sugar sensing in drosophila larvae", *Current Biology*, Vol. 23, No. 15, pp. 1466–1471, 2013.

56. Patat, O., A. Pagin, A. Siegfried, V. Mitchell, N. Chassaing, S. Faguer, L. Monteil, V. Gaston, L. Bujan, M. Courtade-Saïdi, F. Marcelli, G. Lalau, J. M. Rigot, R. Mieusset and E. Bieth, "Truncating Mutations in the Adhesion G Protein-Coupled Receptor G2 Gene ADGRG2 Cause an X-Linked Congenital Bilateral Absence of Vas Deferens", *American Journal of Human Genetics*, Vol. 99, No. 2, pp. 437–442, 2016.

57. Kokot, F. and R. Ficek, "Effects of neuropeptide Y on appetite", *Mineral and Electrolyte Metabolism*, Vol. 25, No. 4-6, pp. 303–305, 1999.

58. Liesch, J., L. L. Bellani and L. B. Vosshall, "Functional and Genetic Characterization of Neuropeptide Y-Like Receptors in Aedes aegypti", *PLoS Neglected Tropical Diseases*, Vol. 7, No. 10, 2013.

59. Watta, H., D. Geetanjali and U. K. Kharmateb, "Somatostatin receptors 1 and 5 heterodimerize with epidermal growth factor receptor: Agonist-dependent modulation of the downstream MAPK signalling pathway in breast cancer cells", *Cellular Signalling*, Vol. 21, No. 3, pp. 428–439, 2009.

APPENDIX A: AMINO ACID ABBREVIATIONS

Amino Acid	**One Letter Abbreviation**	**Three Letter Abbreviation**
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamic acid	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

APPENDIX B: MUTATION CODES

Code of mutation	**Explanation**
AFTPP	p.Ile292Ala substitution
AFTPA	p.[Ile292Ala; Pro296Ala]
AFAPA	p.[Ile292Ala; Thr294Ala; Pro296Ala]
AATPA	p.[Ile292Al; Phe293Ala; Thr294Ala; Pro296Ala]
AFAAA	p.[Ile292Al; Thr294Ala; Pro295Ala; Pro296Ala]
AAAAA	p.[Ile292Al; Phe293Ala; Thr294Ala; Pro295Ala; Pro296Ala]
Ndel	p.1_52del

APPENDIX C: NEUROPEPTIDE QUERIES

Table C.1: The neuropeptide queries used in similarity searches.

Neuropeptide Abbreviation	**Name of Neuropeptide**	***Query Organism***	**Accession Code**
ACP	Adipokinetic hormone/corazonin-like peptide	*Asterias rubens*	ALJ99955.1
		Tribolium castaneum	ADF28807.1
		Rhodnius prolixus	AKO62855.1
		Heliothis virescens	ADW77572.1
AKH	Adipokinetic hormone	*Drosophila melanogaster*	P61855
AMO	Amontillado	*Drosophila melanogaster*	NP_477318.1
AST-A	Allatostatin A	*Apis mellifera*	P85797.1
		Camponotus floridanus	E2ADX8
AST-B	Allatostatin B	*Tribolium castaneum*	NP_001137202.1
		Drosophila melanogaster	Q9VVF7.1
AST-C	Allatostatin C	*Nasonia vitripennis*	ADM26612.1
		Bombyx mori	B3IWA9
		Nilaparvata lugens	U3U451

Table C.1: The neuropeptide queries used in similarity searches (cont.).

Neuropeptide Abbreviation	**Name of Neuropeptide**	***Query Organism***	**Accession Code**
AST-C	Allatostatin C	*Camponotus floridanus*	E2A6Z3
AST-CC	Allatostatin CC	*Nasonia vitripennis*	ADM15719.1
		Chilo suppressalis	A0A0S1U1C0
		Drosophila melanogaster	Q9VKK4
AT	Allatotropin	*Manduca sexta*	AAB08759.1
		Rhodnius prolixus	P85825
BURSA	Bursicon alpha	*Agrilus planipennis*	XP_018318861.1
		Zootermopsis nevadensis	KDR13886.1
		Drosophila melanogaster	AAF55915.1
BURSB	Bursicon beta	*Zootermopsis nevadensis*	KDR13885.1
CAP	Cardioaccelaratory peptide	*Nilaparvata lugens*	BAO00941.1
CCAP	Crustacean cardioactive peptide	*Drosophila melanogaster*	BAO00941.1
CCH	CCHamide-1	*Delia radicum*	B3EWM8

Table C.1: The neuropeptide queries used in similarity searches (cont.).

Neuropeptide Abbreviation	**Name of Neuropeptide**	***Query Organism***	**Accession Code**
CCH	CCHamide-1	*Drosophila melanogaster*	AAF55014.2
CNM	CNMamide	*Drosophila melanogaster*	NP_001189021.1
Corazonin	Corazonin	*Drosophila melanogaster*	AAF55046.1
DH31	Diuretic hormone 31	*Nilaparvata lugens*	U3U8Y9
DH44	Diuretic hormone 44	*Zootermopsis nevadensis*	KDR14744.1
		Periplaneta americana	ALG35940.1
DMM	Dimmed	*Cryptotermes secundus*	XP_023713428.1
EH	Eclosion hormone	*Drosophila melanogaster*	AAF55423.1
ETH	Ecdysis-triggering hormone	*Drosophila melanogaster*	Q9U4J0
FMRF	FMRFamide	*Drosophila melanogaster*	AAF58874.1
FMRF-L	FMRFamide-like protein	*Delia radicum*	B3EWJ9
GHA	Glycoprotein hormone alpha	*Drosophila melanogaster*	NP_001104054.2
GHB	Glycoprotein hormone beta	*Drosophila melanogaster*	NP_001104335.1

Table C.1: The neuropeptide queries used in similarity searches (cont.).

Neuropeptide Abbreviation	**Name of Neuropeptide**	***Query Organism***	**Accession Code**
ILP	Insulin like peptide 2	*Drosophila melanogaster*	AAF50204.1
ITP	Ion transport peptide	*Manduca sexta*	Q1XAU8 and Q1XAU7
Leucokinin	Leucokinin	*Drosophila melanogaster*	NP_524893.2
MS	Myosuppressin	*Drosophila melanogaster*	FBpp0083991, FBpp0306667 and FBpp0312058
NP	Neuroparsin	*Locusta migratoria*	NPAB_LOCMI
NPF	Neuropeptide F	*Drosophila melanogaster*	FBpp0082778, FBpp0304074 and FBpp0304075
NPL	Neuropeptide-like peptide	*Drosophila melanogaster*	FBpp0072348
ORC	Orcokinin	*Homarus americanus*	ACB41787.1
PBAN	Pyrokinin	*Locusta migratoria*	P41488.1 and P85867.1
		Nilaparvata lugens	BAO00974.1
PTTH	Prothoracicotropic hormone protein	*Tribolium castaneum*	AKN79607.1

Table C.1: The neuropeptide queries used in similarity searches (cont.).

Neuropeptide Abbreviation	**Name of Neuropeptide**	***Query Organism***	**Accession Code**
SIF	Neuropeptide SIFamide	*Drosophila melanogaster*	Q6IGX9.1
sNPF	Short neuropeptide F	*Agrilus planipennis*	XP_018330111.1
SK	Sulfakinin	*Drosophila melanogaster*	AAF52173.2
TK	Tachykinin	*Drosophila melanogaster*	AAF54735.1
Trissin	Trissin	*Drosophila melanogaster*	NP_650471.2
V-L	Vasopressin-like	*Tribolium castaneum*	ABX52000.1

APPENDIX D: PCR EFFICIENCIES

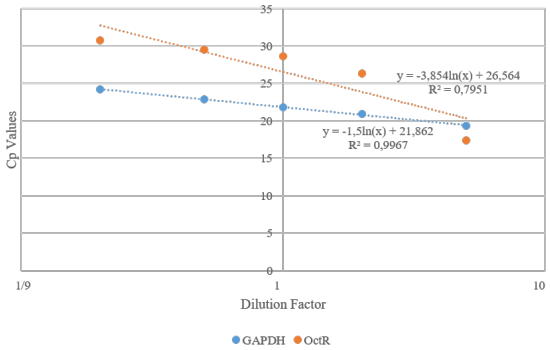


Figure D.1: Samples of GAPDH and CamOctR primer couple standard curves were given. The PCR efficiency values in Table D.1 were obtained from the equations on standard curves.

Table D.1: The table of PCR efficiency for each primer couple in qPCR experiments.

Primer Couple	GAPDH	CamAdgrG2	CamNPYR	CamInoR	CamAlstR-C
PCR Eff.	1,948	1,496847	1,704026	1,462352	1,509623
Primer Couple	Orphan GPCR	CamAlstR-A	CamSPR	CamTyr2R	CamCCKR
PCR Eff.	1,67481	1,468721	1,592388	1,490055	1,207816
Primer Couple	CamCalR	CamAKHR	CamDHR	CamOctR	Gustatory Receptor
PCR Eff.	1,52772	1,443353	1,555389	1,296208	R^2 very bad

APPENDIX E: STATISTICS OF QPCR RESULTS

The statistics of only significant results were given.

Table Analyzed	Data with SD
Two-way ANOVA	Ordinary
Alpha	0,05

Source of Variation	% of total variation	P value	P value summary	Significant?
Interaction	54,75	¡ 0,0001	****	Yes
Row Factor	6,889	¡ 0,0001	****	Yes
Column Factor	18,38	¡ 0,0001	****	Yes

ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Interaction	1433	96	14,93	F (96, 208) = 5,938	P ¡ 0,0001
Row Factor	180,3	8	22,53	F (8, 208) = 8,965	P ¡ 0,0001
Column Factor	481	12	40,08	F (12, 208) = 15,95	P ¡ 0,0001
Residual	522,8	208	2,514		

Number of missing values	26

Number of families	13
Number of comparisons per family	36
Alpha	0,05

Tukey's multiple comparisons test	Mean Diff,	95% CI of diff,	Significant?	Summary
Orphan GPCR				
BCC vs. Ganglia	8,3	4,242 to 12,36	Yes	****
BCC vs. Ov	13,12	9,062 to 17,18	Yes	****
BCC vs. MT	13,82	9,283 to 18,36	Yes	****
BCC vs. C+F	13,43	9,372 to 17,49	Yes	****
BCC vs. GC	13,46	9,402 to 17,52	Yes	****
BCC vs. PH	13,83	9,772 to 17,89	Yes	****
BCC vs. FB	13,61	9,552 to 17,67	Yes	****
BCC vs. Aort	13,43	8,893 to 17,97	Yes	****
Ganglia vs. Ov	4,82	0,7620 to 8,878	Yes	**
Ganglia vs. MT	5,52	0,9830 to 10,06	Yes	**
Ganglia vs. C+F	5,13	1,072 to 9,188	Yes	**
Ganglia vs. GC	5,16	1,102 to 9,218	Yes	**
Ganglia vs. PH	5,53	1,472 to 9,588	Yes	***
Ganglia vs. FB	5,31	1,252 to 9,368	Yes	**
Ganglia vs. Aort	5,13	0,5930 to 9,667	Yes	*
Ov vs. MT	0,7	-3,837 to 5,237	No	ns
Ov vs. C+F	0,31	-3,748 to 4,368	No	ns
Ov vs. GC	0,34	-3,718 to 4,398	No	ns
Ov vs. PH	0,71	-3,348 to 4,768	No	ns
Ov vs. FB	0,49	-3,568 to 4,548	No	ns
Ov vs. Aort	0,31	-4,227 to 4,847	No	ns

MT vs. C+F	-0,39	-4,927 to 4,147	No	ns
MT vs. GC	-0,36	-4,897 to 4,177	No	ns
MT vs. PH	0,01	-4,527 to 4,547	No	ns
MT vs. FB	-0,21	-4,747 to 4,327	No	ns
MT vs. Aort	-0,39	-5,360 to 4,580	No	ns
C+F vs. GC	0,03	-4,028 to 4,088	No	ns
C+F vs. PH	0,4	-3,658 to 4,458	No	ns
C+F vs. FB	0,18	-3,878 to 4,238	No	ns
C+F vs. Aort	0	-4,537 to 4,537	No	ns
GC vs. PH	0,37	-3,688 to 4,428	No	ns
GC vs. FB	0,15	-3,908 to 4,208	No	ns
GC vs. Aort	-0,03	-4,567 to 4,507	No	ns
PH vs. FB	-0,22	-4,278 to 3,838	No	ns
PH vs. Aort	-0,4	-4,937 to 4,137	No	ns
FBvs. Aort	-0,18	-4,717 to 4,357	No	ns

CamInoR				
BCC vs. Ganglia	-0,14	-4,198 to 3,918	No	ns
BCC vs. Ov	-0,29	-4,348 to 3,768	No	ns
BCC vs. MT	-0,08	-4,617 to 4,457	No	ns
BCC vs. C+F	0,03	-4,028 to 4,088	No	ns
BCC vs. GC	-13,87	-17,93 to -9,812	Yes	****
BCC vs. PH	-3,1	-7,158 to 0,9580	No	ns
BCC vs. FB	-0,21	-4,268 to 3,848	No	ns
BCC vs. Aort	-1,7	-6,237 to 2,837	No	ns
Ganglia vs. Ov	-0,15	-4,208 to 3,908	No	ns
Ganglia vs. MT	0,06	-4,477 to 4,597	No	ns
Ganglia vs. C+F	0,17	-3,888 to 4,228	No	ns
Ganglia vs. GC	-13,73	-17,79 to -9,672	Yes	****
Ganglia vs. PH	-2,96	-7,018 to 1,098	No	ns
Ganglia vs. FB	-0,07	-4,128 to 3,988	No	ns

Ganglia vs. Aort	-1,56	-6,097 to 2,977	No	ns
Ov vs. MT	0,21	-4,327 to 4,747	No	ns
Ov vs. C+F	0,32	-3,738 to 4,378	No	ns
Ov vs. GC	-13,58	-17,64 to -9,522	Yes	****
Ov vs. PH	-2,81	-6,868 to 1,248	No	ns
Ov vs. FB	0,08	-3,978 to 4,138	No	ns
Ov vs. Aort	-1,41	-5,947 to 3,127	No	ns
MT vs. C+F	0,11	-4,427 to 4,647	No	ns
MT vs. GC	-13,79	-18,33 to -9,253	Yes	****
MT vs. PH	-3,02	-7,557 to 1,517	No	ns
MT vs. FB	-0,13	-4,667 to 4,407	No	ns
MT vs. Aort	-1,62	-6,590 to 3,350	No	ns
C+F vs. GC	-13,9	-17,96 to -9,842	Yes	****
C+F vs. PH	-3,13	-7,188 to 0,9280	No	ns
C+F vs. FB	-0,24	-4,298 to 3,818	No	ns
C+F vs. Aort	-1,73	-6,267 to 2,807	No	ns
GC vs. PH	10,77	6,712 to 14,83	Yes	****
GC vs. FB	13,66	9,602 to 17,72	Yes	****
GC vs. Aort	12,17	7,633 to 16,71	Yes	****
PH vs. FB	2,89	-1,168 to 6,948	No	ns
PH vs. Aort	1,4	-3,137 to 5,937	No	ns
FBvs. Aort	-1,49	-6,027 to 3,047	No	ns

CamCalR				
BCC vs. Ganglia	-0,84	-4,898 to 3,218	No	ns
BCC vs. Ov	0,83	-3,228 to 4,888	No	ns
BCC vs. MT	0,05	-4,487 to 4,587	No	ns
BCC vs. C+F	0,26	-3,798 to 4,318	No	ns
BCC vs. GC	1,38	-2,678 to 5,438	No	ns
BCC vs. PH	0,51	-3,548 to 4,568	No	ns
BCC vs. FB	-5,99	-10,05 to -1,932	Yes	***

BCC vs. Aort	-1,44	-5,977 to 3,097	No	ns
Ganglia vs. Ov	1,67	-2,388 to 5,728	No	ns
Ganglia vs. MT	0,89	-3,647 to 5,427	No	ns
Ganglia vs. C+F	1,1	-2,958 to 5,158	No	ns
Ganglia vs. GC	2,22	-1,838 to 6,278	No	ns
Ganglia vs. PH	1,35	-2,708 to 5,408	No	ns
Ganglia vs. FB	-5,15	-9,208 to -1,092	Yes	**
Ganglia vs. Aort	-0,6	-5,137 to 3,937	No	ns
Ov vs. MT	-0,78	-5,317 to 3,757	No	ns
Ov vs. C+F	-0,57	-4,628 to 3,488	No	ns
Ov vs. GC	0,55	-3,508 to 4,608	No	ns
Ov vs. PH	-0,32	-4,378 to 3,738	No	ns
Ov vs. FB	-6,82	-10,88 to -2,762	Yes	****
Ov vs. Aort	-2,27	-6,807 to 2,267	No	ns
MT vs. C+F	0,21	-4,327 to 4,747	No	ns
MT vs. GC	1,33	-3,207 to 5,867	No	ns
MT vs. PH	0,46	-4,077 to 4,997	No	ns
MT vs. FB	-6,04	-10,58 to -1,503	Yes	**
MT vs. Aort	-1,49	-6,460 to 3,480	No	ns
C+F vs. GC	1,12	-2,938 to 5,178	No	ns
C+F vs. PH	0,25	-3,808 to 4,308	No	ns
C+F vs. FB	-6,25	-10,31 to -2,192	Yes	****
C+F vs. Aort	-1,7	-6,237 to 2,837	No	ns
GC vs. PH	-0,87	-4,928 to 3,188	No	ns
GC vs. FB	-7,37	-11,43 to -3,312	Yes	****
GC vs. Aort	-2,82	-7,357 to 1,717	No	ns
PH vs. FB	-6,5	-10,56 to -2,442	Yes	****
PH vs. Aort	-1,95	-6,487 to 2,587	No	ns
FBvs. Aort	4,55	0,01304 to 9,087	Yes	*

CamSPR

BCC vs. Ganglia	4,54	0,4820 to 8,598	Yes	*
BCC vs. Ov	12,63	8,572 to 16,69	Yes	****
BCC vs. MT	12,11	7,573 to 16,65	Yes	****
BCC vs. C+F	6,44	2,382 to 10,50	Yes	****
BCC vs. GC	9,87	5,812 to 13,93	Yes	****
BCC vs. PH	11,51	7,452 to 15,57	Yes	****
BCC vs. FB	7,38	3,322 to 11,44	Yes	****
BCC vs. Aort	5,81	1,273 to 10,35	Yes	**
Ganglia vs. Ov	8,09	4,032 to 12,15	Yes	****
Ganglia vs. MT	7,57	3,033 to 12,11	Yes	****
Ganglia vs. C+F	1,9	-2,158 to 5,958	No	ns
Ganglia vs. GC	5,33	1,272 to 9,388	Yes	**
Ganglia vs. PH	6,97	2,912 to 11,03	Yes	****
Ganglia vs. FB	2,84	-1,218 to 6,898	No	ns
Ganglia vs. Aort	1,27	-3,267 to 5,807	No	ns
Ov vs. MT	-0,52	-5,057 to 4,017	No	ns
Ov vs. C+F	-6,19	-10,25 to -2,132	Yes	***
Ov vs. GC	-2,76	-6,818 to 1,298	No	ns
Ov vs. PH	-1,12	-5,178 to 2,938	No	ns
Ov vs. FB	-5,25	-9,308 to -1,192	Yes	**
Ov vs. Aort	-6,82	-11,36 to -2,283	Yes	***
MT vs. C+F	-5,67	-10,21 to -1,133	Yes	**
MT vs. GC	-2,24	-6,777 to 2,297	No	ns
MT vs. PH	-0,6	-5,137 to 3,937	No	ns
MT vs. FB	-4,73	-9,267 to -0,1930	Yes	*
MT vs. Aort	-6,3	-11,27 to -1,330	Yes	**
C+F vs. GC	3,43	-0,6280 to 7,488	No	ns
C+F vs. PH	5,07	1,012 to 9,128	Yes	**
C+F vs. FB	0,94	-3,118 to 4,998	No	ns
C+F vs. Aort	-0,63	-5,167 to 3,907	No	ns
GC vs. PH	1,64	-2,418 to 5,698	No	ns

GC vs. FB	-2,49	-6,548 to 1,568	No	ns
GC vs. Aort	-4,06	-8,597 to 0,4770	No	ns
PH vs. FB	-4,13	-8,188 to -0,07202	Yes	*
PH vs. Aort	-5,7	-10,24 to -1,163	Yes	**
FBvs. Aort	-1,57	-6,107 to 2,967	No	ns

CamTyr2R				
BCC vs. Ganglia	-0,05	-4,108 to 4,008	No	ns
BCC vs. Ov	1,78	-2,278 to 5,838	No	ns
BCC vs. MT	1,88	-2,657 to 6,417	No	ns
BCC vs. C+F	-0,25	-4,308 to 3,808	No	ns
BCC vs. GC	0,62	-3,438 to 4,678	No	ns
BCC vs. PH	-0,85	-4,908 to 3,208	No	ns
BCC vs. FB	-0,29	-4,348 to 3,768	No	ns
BCC vs. Aort	-3,52	-8,057 to 1,017	No	ns
Ganglia vs. Ov	1,83	-2,228 to 5,888	No	ns
Ganglia vs. MT	1,93	-2,607 to 6,467	No	ns
Ganglia vs. C+F	-0,2	-4,258 to 3,858	No	ns
Ganglia vs. GC	0,67	-3,388 to 4,728	No	ns
Ganglia vs. PH	-0,8	-4,858 to 3,258	No	ns
Ganglia vs. FB	-0,24	-4,298 to 3,818	No	ns
Ganglia vs. Aort	-3,47	-8,007 to 1,067	No	ns
Ov vs. MT	0,1	-4,437 to 4,637	No	ns
Ov vs. C+F	-2,03	-6,088 to 2,028	No	ns
Ov vs. GC	-1,16	-5,218 to 2,898	No	ns
Ov vs. PH	-2,63	-6,688 to 1,428	No	ns
Ov vs. FB	-2,07	-6,128 to 1,988	No	ns
Ov vs. Aort	-5,3	-9,837 to -0,7630	Yes	**
MT vs. C+F	-2,13	-6,667 to 2,407	No	ns
MT vs. GC	-1,26	-5,797 to 3,277	No	ns
MT vs. PH	-2,73	-7,267 to 1,807	No	ns

MT vs. FB	-2,17	-6,707 to 2,367	No	ns
MT vs. Aort	-5,4	-10,37 to -0,4300	Yes	*
C+F vs. GC	0,87	-3,188 to 4,928	No	ns
C+F vs. PH	-0,6	-4,658 to 3,458	No	ns
C+F vs. FB	-0,04	-4,098 to 4,018	No	ns
C+F vs. Aort	-3,27	-7,807 to 1,267	No	ns
GC vs. PH	-1,47	-5,528 to 2,588	No	ns
GC vs. FB	-0,91	-4,968 to 3,148	No	ns
GC vs. Aort	-4,14	-8,677 to 0,3970	No	ns
PH vs. FB	0,56	-3,498 to 4,618	No	ns
PH vs. Aort	-2,67	-7,207 to 1,867	No	ns
FBvs. Aort	-3,23	-7,767 to 1,307	No	ns

CamAKHR				
BCC vs. Ganglia	-3,24	-7,298 to 0,8180	No	ns
BCC vs. Ov	1,19	-2,868 to 5,248	No	ns
BCC vs. MT	-0,9	-5,437 to 3,637	No	ns
BCC vs. C+F	0,4	-3,658 to 4,458	No	ns
BCC vs. GC	-0,38	-4,438 to 3,678	No	ns
BCC vs. PH	1,22	-2,838 to 5,278	No	ns
BCC vs. FB	-2,96	-7,018 to 1,098	No	ns
BCC vs. Aort	-2,08	-6,617 to 2,457	No	ns
Ganglia vs. Ov	4,43	0,3720 to 8,488	Yes	*
Ganglia vs. MT	2,34	-2,197 to 6,877	No	ns
Ganglia vs. C+F	3,64	-0,4180 to 7,698	No	ns
Ganglia vs. GC	2,86	-1,198 to 6,918	No	ns
Ganglia vs. PH	4,46	0,4020 to 8,518	Yes	*
Ganglia vs. FB	0,28	-3,778 to 4,338	No	ns
Ganglia vs. Aort	1,16	-3,377 to 5,697	No	ns
Ov vs. MT	-2,09	-6,627 to 2,447	No	ns
Ov vs. C+F	-0,79	-4,848 to 3,268	No	ns

Ov vs. GC	-1,57	-5,628 to 2,488	No	ns
Ov vs. PH	0,03	-4,028 to 4,088	No	ns
Ov vs. FB	-4,15	-8,208 to -0,09202	Yes	*
Ov vs. Aort	-3,27	-7,807 to 1,267	No	ns
MT vs. C+F	1,3	-3,237 to 5,837	No	ns
MT vs. GC	0,52	-4,017 to 5,057	No	ns
MT vs. PH	2,12	-2,417 to 6,657	No	ns
MT vs. FB	-2,06	-6,597 to 2,477	No	ns
MT vs. Aort	-1,18	-6,150 to 3,790	No	ns
C+F vs. GC	-0,78	-4,838 to 3,278	No	ns
C+F vs. PH	0,82	-3,238 to 4,878	No	ns
C+F vs. FB	-3,36	-7,418 to 0,6980	No	ns
C+F vs. Aort	-2,48	-7,017 to 2,057	No	ns
GC vs. PH	1,6	-2,458 to 5,658	No	ns
GC vs. FB	-2,58	-6,638 to 1,478	No	ns
GC vs. Aort	-1,7	-6,237 to 2,837	No	ns
PH vs. FB	-4,18	-8,238 to -0,1220	Yes	*
PH vs. Aort	-3,3	-7,837 to 1,237	No	ns
FBvs. Aort	0,88	-3,657 to 5,417	No	ns

BCC: Brain, Corpora cardiaca and Corpora allata, MT: Malpighian tubules, C: Crop, F: Foregut, GC: Gastric cecea, PH: Post-posterior midgut and Hindgut, FB: Fat Body, Ov: Ovaries