Isolation and Functional Characterization of Calcitonin-Like Diuretic Hormone Receptors in *Rhodnius prolixus*

Meet Zandawala^{1,2*}, Shizhong Li², Frank Hauser², Cornelis J. P. Grimmelikhuijzen², Ian Orchard¹

1 Department of Biology, University of Toronto Mississauga, Mississauga, Ontario, Canada, 2 Center for Functional and Comparative Insect Genomics, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Abstract

Several families of diuretic hormones exist in insects, one of which is the calcitonin-like diuretic hormone (CT/DH) family. CT/DH mediates its effects by binding to family B G-protein coupled receptors (GPCRs). Here we isolate and functionally characterize two R. prolixus CT/DH receptor paralogs (Rhopr-CT/DH-R1 and Rhopr-CT/DH-R2) using a novel heterologous assay utilizing a modified human embryonic kidney 293 (HEK293) cell line. Rhopr-CT/DH-R1 is orthologous to the previously characterized D. melanogaster CT/DH receptor (CG17415) while Rhopr-CT/DH-R2 is orthologous to the D. melanogaster receptor (CG4395), an orphan receptor whose ligand was unknown until now. We determine the cDNA sequences of three splice variants encoding Rhopr-CT/DH-R1 (Rhopr-CT/DH-R1-A, Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C) and two splice variants encoding Rhopr-CT/DH-R2 (Rhopr-CT/DH-R2-A and Rhopr-CT/DH-R2-B). Rhopr-CT/DH-R1-A and Rhopr-CT/DH-R2-A encode truncated receptors that lack six and seven of the characteristic seven transmembrane domains, respectively. Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C, which only differ by 2 amino acids in their C-terminal domain, can both be activated by Rhopr-CT/DH at equal sensitivities (EC₅₀ = 200-300nM). Interestingly, Rhopr-CT/DH-R2-B is much more sensitive to Rhopr-CT/DH (EC₅₀ = 15nM) compared to Rhopr-CT/DH-R1-B/C and also yields a much greater response (amplitude) in our heterologous assay. This is the first study to reveal that insects possess at least two CT/DH receptors, which may be functionally different. Quantitative PCR demonstrates that Rhopr-CT/DH-R1 and Rhopr-CT/DH-R2 have distinct expression patterns, with both receptors expressed centrally and peripherally. Moreover, the expression analysis also identified novel target tissues for this neuropeptide, including testes, ovaries and prothoracic glands, suggesting a possible role for Rhopr-CT/DH in reproductive physiology and development.

Citation: Zandawala M, Li S, Hauser F, Grimmelikhuijzen CJP, Orchard I (2013) Isolation and Functional Characterization of Calcitonin-Like Diuretic Hormone Receptors in *Rhodnius prolixus*. PLoS ONE 8(11): e82466. doi:10.1371/journal.pone.0082466

Editor: Hubert Vaudry, University of Rouen, France, France

Received September 24, 2013; Accepted October 23, 2013; Published November 29, 2013

Copyright: © 2013 Zandawala et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant to I.O., an NSERC Canadian Graduate Scholarship and NSERC Michael Smith Foreign Study Supplement to M.Z., and by grants from the Danish Research Agency, Carlsberg Foundation, and Lundbeck Foundation to C.J.P.G. and F.H. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

* E-mail: meet.zandawala@utoronto.ca

Introduction

Various neurohormone families have been implicated in regulating diuresis in insects. One such family is the calcitoninlike diuretic hormone (CT/DH) family which is related to the mammalian calcitonin and calcitonin gene-related peptide hormonal system. The first member of this peptide family in insects was isolated and functionally characterized in *Diploptera punctata* [1]. This peptide was originally termed diuretic hormone 31 (DH₃₁) due to its ability to stimulate Malpighian tubule (MT) secretion in certain insects and due to the fact that it is comprised of 31 amino acids [1-3]. As is the case with many peptides that are named because of a particular bioassay involved in their isolation, regulating diuresis may not be their function in other insects. Thus, CT/DHs do not stimulate MT secretion in *Acrosternum hilare* and *Podisus maculiventris* [4]. Moreover, the role of *Rhodnius prolixus* CT/DH (Rhopr-CT/DH) in diuresis is also questionable as it does not stimulate water reabsorption across the midgut, and only stimulates MT secretion at a rate which is 1.5% of maximum [5,6]; however, it may play a broad role in feedingrelated physiological events in various insects. For example, Rhopr-CT/DH has been shown to have myostimulatory effects on hindgut, dorsal vessel and salivary glands whereas the *Drosophila melanogaster* CT/DH is required for peristalsis in the larval midgut [7-9]. Furthermore, *D. punctata* CT/DH analogs have anorexigenic effects in *Locusta migratoria* nymphs [10]. It is thus evident that CT/DHs, like several other neuropeptides, are pleiotropic in nature. Hence, in order to elucidate additional physiological roles for these hormones, it is important to identify and characterize their receptors and determine their expression patterns.

Insect CT/DH receptors (CT/DH-Rs) belong to the family of secretin-like (family B 1) G-protein coupled receptors (GPCRs) [11]. Johnson et al. characterized the first insect CT/DH-R from D. melanogaster in 2005 [12]. Signaling through this receptor was shown to be dependent on accessory proteins (receptor activity modifying proteins (RAMPs) and receptor component protein (RCP)), in a manner analogous to mammals [13,14]. Recently, a receptor orthologous to this was functionally characterized in Aedes aegypti (AedaeGPCRCAL1) via RNAibased knockdown [15,16]. RNAi treated females showed a 30% reduction in fluid excretion (relative to control groups) following a blood meal and the hindguts exhibited a 50% reduction in contraction frequency in response to A. aegypti CT/DH (Aedae-CT/DH) compared to controls. Moreover, a 57% decrease in fluid secretion in response to Aedae-CT/DH was also observed in MTs in which AedaeGPCRCAL1 was knocked-down [15].

In the present study, we have isolated and characterized a CT/DH-R from R. prolixus that is orthologous to the previously characterized CT/DH-Rs in D. melanogaster and A. aegypti [12,15]. We propose to rename these receptors as CT/DH-R1. Moreover we have also isolated and characterized another family B1 GPCR from R. prolixus that is orthologous to the D. melanogaster receptor (CG4395), hector. This orphan receptor is also related to insect CT/DH-Rs and is activated by Rhopr-CT/DH. Hence, we propose to name these receptors as CT/DH-R2. Rhopr-CT/DH-R2 is much more sensitive to Rhopr-CT/DH compared to Rhopr-CT/DH-R1 in our heterologous assay utilizing human embryonic kidney (HEK)-293 cells stably expressing a modified cyclic nucleotide-gated (CNG) channel (HEK293/CNG). We obtained robust and sensitive responses in these cells without having to co-express any accessory proteins, making it ideal to study CT/DH-Rs and, perhaps, deorphanize other family B1 GPCRs. To our knowledge, this is the first study to reveal that insects possess at least two CT/DH receptors, which may be functionally different. Quantitative PCR demonstrates that Rhopr-CT/DH-R1 and Rhopr-CT/DH-R2 have distinct expression patterns, with both receptors expressed centrally and peripherally.

Materials and Methods

Animals

Fifth-instar and adult *R. prolixus* (4-5 weeks post-feeding) were raised in a long standing colony that was maintained in incubators at 60% humidity and 25°C. The insects were routinely fed artificially once in each instar on defibrinated rabbit blood (Hemostat Laboratories, Dixon, CA, USA; supplied by Cedarlane Laboratories Inc., Burlington, ON, Canada).

Isolation of cDNA sequences encoding R. prolixus CT/DH receptors

Supercontigs in FASTA format, representing the *R. prolixus* preliminary genome assembly (June 2009 release), were downloaded from the genome server at The Genome Institute at Washington University (<u>http://genome.wustl.edu/pub/</u>

organism/Invertebrates/Rhodnius_prolixus/). These supercontigs were then imported into Geneious Pro 4.7.6 and used to perform local tBLASTn search, with the D. melanogaster CT/DH receptor (CG17415, accession no: NP 725278.1) protein sequence acting as the query. Hits along two different supercontigs were obtained; these represent two putative CT/DH receptors. Primers specific to the hit regions were designed (Table S1 in File S1) and used to amplify the partial cDNA sequence encoding Rhopr-CT/DH-R1 and Rhopr-CT/DH-R2. Template for the PCR was cDNA synthesized using total RNA extracted from individuallydissected tissues (see section: Quantitative PCR tissue profiling). PCR was performed using s1000 thermal cycler (Bio-Rad Laboratories, Mississauga, ON, Canada) with a temperature-cycling profile that consisted of an initial denaturation (94°C for 3 min) and 35 cycles of denaturation (94°C for 30 sec), annealing (59°C for 30 sec) and extension (72°C for 1 min); a final 10 min extension at 72°C was also included. Gel electrophoresis was used to visualize the PCR product which was then extracted using the EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Inc., Markham, ON, Canada). The gel extracted product was cloned and sequenced using the methods described earlier [17].

Complete cDNA sequences encoding the two receptors were obtained using a modified 5' and 3' rapid amplification of cDNA ends (RACE) PCR technique, as described earlier [17]. Primers used for 5' and 3' RACE PCRs have been listed in Table S2 and Table S3 in File S1, respectively. Lastly, the largest cDNA fragments encoding the receptors were amplified using the primers listed in Table S4 in File S1 and a proofreading Taq polymerase. The PCR products were cloned and sequenced as explained earlier [17].

Sequence analysis

The intron-exon boundaries were predicted using a combination of a BLAST search of the R. prolixus genome and Genie, an online software for splice site prediction [18]. Membrane topology of the receptors was predicted using the Transmembrane Prediction Tool plugin for Geneious. The potential phosphorylation sites were predicted using the NetPhos 2.0 Server [19] and the potential N-linked glycosylation sites predicted using the NetNGlyc 1.0 Server. Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/ - last accessed on August 1, 2013) was used to align Rhopr-CT/DH-R1 (KC660148, KC660149 and KC660150) and R2 isoforms (KF446640 and KF494337) with its homologs from D. melanogaster (NP_725278.1 and NP_572843.2) and Aedes aegypti (AEU12191.1). The alignment figure was obtained using the BOXSHADE 3.21 server (http://www.ch.embnet.org/ software/BOX form.html - last accessed on August 1, 2013).

Additional family B1 GPCR amino acid sequences were included for phylogenetic analysis. These included corticotropin releasing-factor (CRF)-related diuretic hormone (CRF/DH) receptors (CRF/DH-Rs), pigment dispersing factor (PDF) receptors (PDF-Rs) and CT/DH-Rs from a variety of insects. Moreover, CRF receptors (NP_001138618.1 and NP_001189404.1), calcitonin receptor (CTR) (NP_001158209.1) and calcitonin receptor-like receptor

(CRLR) (NP_001258680.1) from *Homo sapiens* were also included in the analysis and *D. melanogaster* metabotropic glutamate receptor (NP_524639.2) was utilized as an outgroup. ClustalX2 was used to align these sequences and the alignment exported to MEGA5 [20,21]. A maximum parsimonious tree was constructed using Close-Neighbor-Interchange (CNI) analysis and the bootstrap values obtained were based on 1000 replicates.

Preparation of expression vectors

The largest cDNA fragments encoding *Rhopr-CT/DH-R1* transcript variants and *Rhopr-CT/DH-R2-B* were amplified as described earlier (see section: Isolation of cDNA sequences encoding *R. prolixus* CT/DH receptors). The PCR products from these reactions were used as a template to amplify the ORF and introduce a Kozak translation initiation sequence at the 5' end using the primers listed in Table S5 in File S1. The resulting products were cloned into pGEM-T Easy vector (Promega, Madison, WI, USA). These were then subcloned into either pIRES2-ZsGreen1 (Clontech, Mountain View, CA, USA) or pcDNA 3.1⁺ (Life Technologies Corporation, Carlsbad, CA, USA) for expression in mammalian cells.

Cell culture and transfections

Human embryonic kidney (HEK)-293 cells stably expressing a modified cyclic nucleotide-gated (CNG) channel (HEK293/ CNG) (previously available through BD Biosciences, Mississauga, ON, Canada) were used to functionally characterize the receptors [22]. HEK293/CNG cells were grown in Dulbecco's Modified Eagle Medium Nutrient Mixture F12-Ham (DMEM/F12) and supplemented with 10% heatinactivated fetal bovine serum (FBS), 1% penicillin and streptomycin and 250µg/mL G418 (Life Technologies Corporation, Carlsbad, CA, USA). The cells were incubated at 37°C in 5% CO₂. X-tremeGENE HP DNA transfection reagent (Roche Applied Science, Indianapolis, IN, USA) was used to transiently co-transfect the cells with the expression vectors containing receptor transcript variant and cytoplasmic luminescent reporter aequorin at ratio of 2:1 (transfection reagent to expression vectors) using the manufacturer recommended protocol. For negative control, empty expression vector without any receptor transcript was also used to transfect the cells. Cells were incubated for 48 hours and then used to perform the bioluminescence assay.

Alternatively spliced transcript variants of mammalian calcitonin receptors have been known to dimerize with normal functional receptors and inhibit either their surface expression or ligand-induced intracellular cAMP production [23,24]. To determine if Rhopr-CT/DH-R1-A interacted with Rhopr-CT/DH-R1-B/C, we utilized Chinese hamster ovary (CHO) cells stably expressing the human G-protein G16 (CHO/G16) [25]. CHO/G16 cells stably expressing either Rhopr-CT/DH-R1-B or Rhopr-CT/DH-R1-C were grown and transfected as described earlier [26]. These cells were then transiently co-transfected with Rhopr-CT/DH-R1-A and aequorin [26].

Bioluminescence assay

Bioluminescence assay using CHO/G16 cells was performed as described previously [25-27]. To perform the assay using HEK293/CNG cells, they were first harvested by incubating in a PBS-EDTA solution and resuspended in bovine serum albumin (BSA) media (DMEM/F12 containing 1% BSA and 1% penicillin and streptomycin). Coelenterazine h (Promega, Madison, WI, USA) was then added to the cells at a 5µM final concentration and incubated for 3 hours with stirring in the dark. The cells were then diluted 10-fold using BSA media and used to perform the assay. Various doses of peptides were prepared in BSA media and plated in triplicates across a 96-well plate. Cells were loaded in each well using an automated injector and the luminescence recorded over 20 seconds using a Wallac Victor2 plate reader (Perkin Elmer, San Diego, CA, USA). Rhopr-CT/DH

(GLDLGLSRGFSGSQAAKHLMGLAAANYAGGPamide) and Rhopr-CRF/DH

(MQRPQGPSLSVANPIEVLRSRLLLEIARRRMKEQDASRVSK NRQYLQQIamide) used for the assay was custom synthesised by GenScript (Piscataway, NJ, USA) at > 95% purity. *D. melanogaster* PDF (NSELINSLLSLPKNMNDAamide) was custom synthesized by GeneMed Synthesis (San Antonio, TX, USA) at > 95% purity. Dose-response curves were obtained and the EC₅₀ values determined using Prism5 software.

Quantitative PCR tissue profiling

The following tissues were individually-dissected from fifthinstar R. prolixus of both sexes and used for spatial expression analysis: (1) CNS, (2) dorsal vessel, (3) fat body, abdominal nerves, diaphragms and trachea, (4) foregut, (5) salivary glands, (6) anterior midgut, (7) posterior midgut, (8) MTs, (9) hindgut, (10) immature testes, (11) immature ovaries and (12) prothoracic glands. The following tissues were dissected from adult R. prolixus to determine the expression pattern in reproductive tissues: (1) testes (2) rest of the male reproductive tissues (3) ovaries and (4) rest of the female reproductive tissues. Total RNA was isolated from these tissues using PureLink® RNA Mini Kit (Life Technologies Corporation, Carlsbad, CA, USA) which was then used to synthesize cDNA using iScript[™] Reverse Transcription Supermix for RT-gPCR (Bio-Rad Laboratories Ltd., Mississauga, ON, Canada). This cDNA was diluted 10-fold and subsequently used as a template for the qPCR reaction. Primers specific for each receptor variant were designed over exon-exon boundaries to determine expression levels for each transcript (Table S6 in File S1). Since the difference between Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C cDNA sequences was minor (see section: Rhopr-CT/DH receptors), primers differentiating these two transcripts could not be designed. The primer efficiencies for each target were calculated and delta-delta Ct method was used to determine the relative expression of each transcript. Geometric averaging of the transcript levels of three housekeeping genes (alpha-tubulin, beta-acting and ribosomal protein 49) was used to normalize the expression levels of the receptor transcripts. Experiments were performed using MX4000 Quantitative PCR System (Stratagene, Mississauga, ON, Canada) with a temperature-cycling profile that consisted of an initial

denaturation (95°C for 30 sec) and 40 cycles of denaturation (95°C for 5 sec) and annealing/extension (60°C for 24 sec); this was followed by a melt curve analysis (60°C - 95°C). SsoFastTM EvaGreen® Supermix with Low ROX (Bio-Rad Laboratories Ltd., Mississauga, ON, Canada) was used to perform all experiments, which included a no template control and 2 technical replicates per reaction. Reactions for each target were run on a gel to confirm the amplicon size. The products were also verified by sequencing.

Results

Rhopr-CT/DH receptors

We have isolated cDNA sequences for three splice variants encoding Rhopr-CT/DH-R1 (Rhopr-CT/DH-R1-A, Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C) and two splice variants encoding Rhopr-CT/DH-R2 (Rhopr-CT/DH-R2-A and Rhopr-CT/DH-R2-B). Rhopr-CT/DH-R1-A, Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C are 1746, 1664 and 1301 nucleotides long and encode receptors comprising of 143, 411 and 409 amino acids, respectively (Figure S1 in File S1 and Figure 1A). The untranslated regions for Rhopr-CT/DH-R1-C could not be cloned via RACE PCRs as primers differentiating between Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C could not be designed. All three variants contain a polyadenylation signal sequence in their 3' UTR. Rhopr-CT/DH-R1-B, Rhopr-CT/DH-R1-C and Rhopr-CT/DH-R2-B all contain seven transmembrane domains, an extracellular N-terminus and an intracellular C-terminus, typical of all GPCRs (Figure 1A, B). They also contain 6 highly-conserved cysteine residues (typical of family B1 GPCRs) and 3 predicted N-linked glycosylation sites in their N-terminus, and various predicted phosphorylation sites in their intracellular domains. Rhopr-CT/DH-R1-A is a truncated version of the receptor and only contains the extracellular N-terminus and a single transmembrane domain (Figure S1 in File S1). The gene encoding this truncated receptor comprises of 14 exons that are separated by 13 introns (Figure 2A). Rhopr-CT/DH-R1-A contains exon 8 that is absent in the other two variants and results in a premature stop codon. The ORF for this variant spans across exons 4 to 8. Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C differ by only 6 nucleotides within their ORF, which results in a 2 amino acid difference between these variants within their intracellular Cterminal domain. Rhopr-CT/DH-R1-C utilizes an alternate splice site in exon 13 which results in that exon being 6 nucleotides shorter at the 3' end. The ORF for these two variants spans across exons 4 to 14.

Rhopr-CT/DH-R2-A and *Rhopr-CT/DH-R2-B* are 1146 and 1639 nucleotides long and yield proteins comprising of 122 and 410 amino acids, respectively (Figure S2 in File S1 and Figure 1B). Rhopr-CT/DH-R2-A comprises a partial N-terminus (contains only 5 cysteine residues) and lacks the seven transmembrane domains (Figure S2 in File S1). Rhopr-CT/DH-R2-B, on the other hand, contains all the characteristics of a family B1 GPCR (Figure 1B). Moreover, it contains 7 predicted N-linked glycosylation sites and one predicted phosphorylation site. The gene encoding these 2 receptor variants is made up of 9 exons (Figure 2B). Exon 4 is absent in *Rhopr-CT/DH-R2-A*

which results in a frame shift and truncated ORF. The ORF for this variant spans across exons 1 to 6 whereas the one for *Rhopr-CT/DH-R2-B* spans across all 9 exons. The untranslated regions for *Rhopr-CT/DH-R2-A* could not be cloned due to its low expression.

Functional receptor assay

To confirm that Rhopr-CT/DH is the ligand for the isolated putative Rhopr-CT/DH-R1 and Rhopr-CT/DH-R2, we expressed these receptors in HEK293/CNG and monitored ligand-receptor interaction using a calcium mobilization assay. Only the receptor isoforms which contained all the 7 transmembrane domains were used in this assay. Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C, which only differ by 2 amino acids in the C-terminus, were both activated by Rhopr-CT/DH with EC_{50} values ranging from 150-300nM (Figure 3A). The maximum luminescence response obtained for both these receptors following the addition of Rhopr-CT/DH was at least 42-fold higher compared to the addition of medium alone. Interestingly, Rhopr-CT/DH-R2-B is much more sensitive to Rhopr-CT/DH (EC₅₀ = 15nM) compared to Rhopr-CT/DH-R1-B/C and results in a greater response (191-fold higher than basal response) in our heterologous assay (Figure 3B). None of these receptors were activated by Rhopr-CRF or Drome-PDF (data not shown). Moreover, no response was observed following the addition of Rhopr-CT/DH to the cells that were transfected with empty vector.

Transfection of Rhopr-CT/DH-R1-A in CHO/G16 cells stablyexpressing either Rhopr-CT/DH-R1-B or Rhopr-CT/DH-R1-C did not influence their sensitivity or kinetics of the response following the addition of Rhopr-CT/DH (data not shown). Since we had stably expressed Rhopr-CT/DH-R1-B in CHO/G16 cells we also compared the kinetics of the response in these cells with that of the HEK293/CNG cells. Rhopr-CT/DH produced a rapid response, with the peak response for HEK/CNG cells and CHO/G16 cells between 5-10 seconds and 0-5 seconds, respectively (Figure S3 in File S1).

Sequence and phylogenetic analysis

Rhopr-CT/DH-R amino acid sequences were aligned along with those of Drome-CT/DH-R1, Aedae-CT/DH-R1 (previously referred to as *Aedae*GPCRCAL1) and Drome-CT/DH-R2 (previously an orphan and also referred to as *hector*). The multiple sequence alignment illustrates high conservation across both the receptors (Figure 4). This conservation is localized not only over the seven transmembrane domains but also over the N'-terminal extracellular domain. Positions of the two predicted N-linked glycosylation sites (positions 85 and 100 in Rhopr-CT/DH-R1-A) and the 6 cysteine residues are conserved across most sequences. Since the N'-terminus forms part of the ligand-binding domain, it is not surprising that CT/DH activates both the receptors [28,29].

Phylogenetic analysis of family B1 GPCRs reveals three main monophyletic groups (Figure 5). These groups represent the three main receptor types – CRF/DH-Rs, PDF-Rs and CT/DH-Rs. All insect CT/DH-Rs are sister to human CTR and CRLR. This further supports the suggestion that these hormonal systems are evolutionary related. Within the clade

A	AAA ATG	ATAA GACT	GTTC' GAAG'	TGAT. TAGT	ACAA' GTGA(TACT' CAAG'	IGTT. IGC T (ATTC GAGC.	GTTT(AGAT'	GCAA TGAT.	CCAG ATTC	AGAA' CAGT	TCAT GATG	CAAG GCGT	CACA	GAAA' GACA(TTAA CTCT	GACA) ITGC(AAGG CACT.	GACTI AGATI	AAGG/ AATA1	AGAAJ FTTT(ATGT(GATG(5' - GCCT <i>I</i> CTAA(GAA AAGA CTCA	GTTG GATC ATTG	TGCA TGTC ACAT	AAGT GACT AGCT	TTGT(GTTT: GTTG <i>l</i>	GGAG ITTA AACG	23 142 261
	ATG MET	TCG Ser	GAT Asp	GAA Glu	ACA Thr	GGC Gly	AAC Asn	CAA Gln	TCA Ser	TTT Phe	CTG Leu	GAC Asp	CCA Pro	CAT His	GCT Ala	GAA Glu	CTT Leu	GTA Val	AAT Asn	TCA Ser	CGC Arg	TAC Tyr	CTT Leu	CAA Gln	TGT Cys	CTA Leu	ACA Thr	ACC Thr	ATC Ile	AAT Asn	351 30
	GAG Glu	TCT Ser	TTG Leu	TCC Ser	AGG Arg	TCT Ser	CTT Leu	CAA Gln	GGA Gly	CTG Leu	CAG Gln	TGT Cys	GAA Glu	GCT Ala	ACT Thr	TTT Phe	GAT Asp	GGA Gly	TGG Trp	TCT Ser	TGC <mark>Cys</mark>	TGG Trp	CCG Pro	GCC Ala	ACA Thr	TCT Ser	GCA Ala	GGA Gly	GAA Glu	ACA Thr	441 60
	GCC Ala	TTC Phe	GCC Ala	CCA Pro	TGT <mark>Cys</mark>	CCA Pro	CAT His	TTC Phe	ATC Ile	ACC Thr	GGC Gly	TTT Phe	GAT Asp	CCA Pro	AAT Asn	CGA Arg	TTG Leu	GCG Ala	CAT His	AAA Lys	GAG Glu	TGC <mark>Cys</mark>	ACA Thr	GAA Glu	AAT Asn	GGT Gly	ACA Thr	TGG Trp	TTT Phe	AGA Arg	531 90
	CAT His	CCT Pro	GAA Glu	TCT Ser	GGA Gly	CAA Gln	ATT Ile	TGG Trp	TCT Ser	AAT Asn	TAT Tyr	ACA Thr	ACA Thr	TGT <mark>Cys</mark>	GTA Val	AAT Asn	TTG Leu	GAT Asp	GAT Asp	TTA Leu	AAT Asn	T TA Leu	AGG Arg	CAA Gln	CAG Gln	GTA Val	AAT Asn	AAC Asn	ATA Ile	TAT Tyr	621 120
	CAG Gln	GCG Ala	GGT Gly	TAC Tyr	TTC Phe	ATA Ile	TCG Ser	CTC Leu	TTA Leu	GCT Ala	CTA Leu	CTC Leu	CTG Leu	TCT Ser	CTC Leu	TTC Phe	ATC Ile	CTA Leu	TCT Ser	TAT Tyr	TTT Phe	AAA Lys	TCT Ser	TTA Leu	AGA Arg	TGT Cys	CCT Pro	CGC Arg	AAT Asn	ACA Thr	711 150
	CTT Leu	CAC His	ATG Met	AAC Asn	TTA Leu	TTT Phe	ACG Thr	GCC Ala	TTT Phe	GCC Ala	TTT Phe	AAC Asn	AAT Asn	TTT Phe	CTG Leu	TGG Trp	CTT Leu	CTC Leu	TGG Trp	TAT Tyr	AGG Arg	CTG Leu	GTC Val	ATT Ile	CCA Pro	TTT Phe	CCG Pro	GAG Glu	GTA Val	ATT Ile	801 180
	CTG Leu	GAA Glu	AAC Asn	GG T Gly	GTA Val	TGG Trp	TGC Cys	CAA Gln	TGT Cys	TTA Leu	CAT His	GTG Val	ATT Ile	CTC Leu	CAC His	TAT Tyr	TTC Phe	TTG Leu	TTA Leu	AGC Ser	TGC Cys	TAT Tyr	GCA Ala	TGG Trp	ATG Met	CTC Leu	GCA Ala	GAA Glu	GGT Gly	GTA Val	891 210
	TAT Tyr	TTA Leu	CAT His	ACT Thr	TTA Leu	TTA Leu	GTG Val	TCT Ser	GCA Ala	TTT Phe	ACC Thr	AGC Ser	GAA Glu	CAG Gln	AAA Lys	TTG Leu	GTT Val	AAA Lys	GTG Val	CTC Leu	ACA Thr	GTA Val	CTT Leu	TCG Ser	TGG Trp	TTC Phe	GTC Val	CCT Pro	ATA Ile	GTC Val	981 240
	TTT Phe	ATC Ile	ACA Thr	TTG Leu	TAT Tyr	ACA Thr	ACG Thr	TTG Leu	AGG Arg	TTA Leu	GCC Ala	TCC Ser	GGC Gly	CAT His	ACT Thr	GAC Asp	C <mark>AA</mark> Gln	TGC Cys	TGG Trp	ATA Ile	GAC Asp	GAA Glu	AGC Ser	GAT Asp	TCA Ser	AAT Asn	ACA Thr	GTA Val	CTT Leu	ATA Ile	1071 270
	ATA Ile	TTG Leu	GTT Val	GCA Ala	ACA Thr	TCA Ser	ATG Met	GGA Gly	CTA Leu	AAC Asn	TTT Phe	ATC Ile	TTC Phe	TTG Leu	TGC Cys	AAC Asn	ATA Ile	ATG Met	CGA Arg	GTG Val	GTT Val	GTT Val	GGT Gly	AAG Lys	TTA Leu	AGA Arg	GCT Ala	GGC Gly	CCA Pro	GCT Ala	1161 300
	CAG Gln	TCT Ser	TCC Ser	AGA Arg	CCT Pro	TCA Ser	AGA Arg	GCC Ala	TTA Leu	CTA Leu	CAA Gln	GCA Ala	CTC Leu	AGA Arq	GCA Ala	ACT Thr	TTA Leu	TTA Leu	TTG Leu	TTA Leu	CCA Pro	TTG Leu	CTT Leu	GGG Gly	CTT Leu	AAC Asn	TAT Tyr	CTT Leu	TTG Leu	ACT Thr	1251 330
	CCA Pro	TTT Phe	CGC Arg	CCA Pro	CCA Pro	AAC Asn	AAT Asn	CAT His	CCG Pro	TGG Trp	GAA Glu	ACT Thr	TAT Tyr	TAT Tyr	GAG Glu	CTA Leu	ATA Ile	TCT Ser	GCT Ala	GTT Val	ACA Thr	GCG Ala	TCC Ser	TTC Phe	CAA Gln	GGT Gly	TTA Leu	TGT Cys	GTT Val	GCT Ala	1341 360
	ACG Thr	CTC Leu	TTC Phe	TGT Cys	TTT Phe	TGT Cys	AAT Asn	GGA Gly	GAG Glu	GTA Val	ATA Ile	GCC Ala	CAA Gln	ATC Ile	AAA Lys	AGG Arg	AAA Lys	TGG Trp	CAG Gln	TAC Tyr	GCT Ala	ATG Met	TTC Phe	AGG Arg	ACT Thr	AGA Arg	GCG Ala	AAC Asn	TCT Ser	TAC Tyr	1431 390
	ACA Thr	GCA Ala	ACC Thr	ACG Thr	GTA Val	TCA Ser	T TT Phe	GTA Val	AGA Arg	TCT Ser	AAT Asn	GCG Ala	GCT Ala	CCT Pro	GTT Val	GCT Ala	GAA Glu	GAA Glu	GAA Glu	AAT Asn	GTT Val	TAA *	TTT	CTTC	TTAT.	AAAG.	AAAG.	AATT	AATA/	<u>AA</u> AC	1528 411
	TTA CCA	FTGA FTAG	AATG ATGT	GGTC TTCA	tcaa ca -	ТАТА З'	AAAT	AATG	TAGA	TAAA.	AAAT	TGTA.	AATA	TGTA	FAAT.	ACTTI	ACG	IGAA!	TTTG'	IGTT	TGAG	ATAT	ATGT	ATTA	FATT.	AATT	GTAA	ATTT.	AGTT:	ITGT	$1647 \\ 1664$
В	5'	- AC	CACC	TCCG	AGTA	CACT	AGTG.	ACGC.	AACA	AAAC.	AGAA	ATTT.	ACGG	CTCC	CTTC	IGTA	CCTG	AATA	TAAG	CACC	AATC	TAAA	CTCC	GCGA	ICGA.	AATT	TTTT	AAAA	AAAG	GAAC	114
В	5' ATG MET	- AC GGA Gly	CACC' AAT Asn	TCCG GTA Val	AGTA GAC Asp	CACTI AAT Asn	AGTG. <u>TTA</u> Leu	ACGC. ACG Thr	AACA ACA Thr	AAAC. CAA Gln	AGAA TCA Ser	ATTT. AAT Asn	ACGG CTA Leu	CTCC CAC His	CTTC TCC Ser	IGTAC AAC Asn	CCTGA ATC Ile	AATA' CTA Leu	TAAG AGA Arg	CACC. TAT Tyr	AATC CTG Leu	FAAA AAA Lys	CTCC GGA Gly	GCGA TTA Leu	ICGA CAA Gln	AATT AGG Arg	TTTT GAA Glu	AAAA TGT <mark>Cys</mark>	AAAGO GAC Asp	GAAC TTA Leu	114 204 30
В	5' ATG MET AGG Arg	- AC GCA Gly AAG Lys	CACC AAT Asn AGA Arg	TCCG GTA Val GCT Ala	AGTA GAC Asp CAA Gln	CACT AAT Asn TTT Phe	AGTG <u>TTA</u> Leu CAA Gln	ACGC. ACG Thr TAT Tyr	AACA ACA Thr TTA Leu	AAAC. CAA Gln TCG Ser	AGAA TCA Ser ACC Thr	ATTT. AAT Asn ATT Ile	ACGG CTA Leu TTG Leu	CTCC CAC His CCG Pro	CTTC TCC Ser <mark>AA</mark> T Asn	IGTAC AAC Asn <u>GTA</u> Val	CCTG ATC Ile AGC Ser	AATA' CTA Leu GAT Asp	TAAG AGA Arg CTA Leu	CACC TAT Tyr AAA Lys	AATC CTG Leu GTG Val	TAAA AAA Lys TAC Tyr	CTCC GGA Gly TGC <mark>Cys</mark>	GCGA TTA Leu CCA Pro	ICGA CAA Gln GCG Ala	AATT AGG Arg ACA Thr	TTTT GAA Glu TTC Phe	AAAA TGT <mark>Cys</mark> GAC Asp	AAAGO GAC Asp GGC Gly	GAAC TTA Leu TGG Trp	114 204 30 294 60
В	5' ATG MET AGG Arg TCA Ser	- AC GGA Gly AAG Lys TGT Cys	CACC AAT Asn AGA Arg TGG Trp	TCCG GTA Val GCT Ala AAT Asn	AGTA GAC Asp CAA Gln <u>ACT</u> Thr	CACTI AAT Asn TTT Phe ACG Thr	AGTG TTA Leu CAA Gln CCA Pro	ACGC. ACG Thr TAT Tyr TCC Ser	AACA ACA Thr TTA Leu GGT Gly	AAAC. CAA Gln TCG Ser GAA Glu	AGAA TCA Ser ACC Thr ATC Ile	ATTT. AAT Asn ATT Ile GCT Ala	ACGG CTA Leu TTG Leu TTG Leu	CTCC CAC His CCG Pro GCA Ala	CTTC TCC Ser AAT Asn CCG Pro	IGTAC AAC Asn G <u>TA</u> Val TGT Cys	CCTG ATC Ile AGC Ser CCA Pro	AATA CTA Leu GAT Asp AAC Asn	IAAG AGA Arg CTA Leu TTT Phe	CACC TAT Tyr AAA Lys GTC Val	AATC CTG Leu GTG Val ACA Thr	TAAA AAA Lys TAC Tyr GGA Gly	CTCC GGA Gly TGC Cys TTT Phe	GCGA TTA Leu CCA Pro GAT Asp	ICGA CAA Gln GCG Ala ATA Ile	AATT AGG Arg ACA Thr AAT Asn	GAA Glu TTC Phe AG G Arg	AAAA TGT <mark>Dys</mark> GAC Asp TTT Phe	AAAGO GAC Asp GGC Gly GCT Ala	GAAC TTA Leu TGG Trp TTT Phe	114 204 30 294 60 384 90
В	5' ATG MET AGG Arg TCA Ser CGC Arg	- AC GGA Gly AAG Lys TGT Cys AAA Lys	CACC AAT Asn AGA Arg TGG Trp TGT Cys	TCCG GTA Val GCT Ala <u>AAT</u> Asn TTA Leu	AGTA GAC Asp CAA Gln ACT Thr GAA Glu	CACT AAT Asn TTT Phe ACG Thr AAT ASN	AGTG TTA Leu CAA Gln CCA Pro GGT Gly	ACGC. Thr TAT Tyr TCC Ser ACT Thr	AACA Thr TTA Leu GGT Gly TGG Trp	AAAC. Gln TCG Ser GAA Glu TTT Phe	AGAA Ser ACC Thr ATC Ile CGA Arg	ATTT. AAT Asn ATT Ile GCT Ala CAT His	ACGG CTA Leu TTG Leu TTG Leu CCA Pro	CTCC His CCG Pro GCA Ala GAC Asp	CTTC Ser AAT Asn CCG Pro ACT Thr	IGTAC AAC Asn GTA Val TGT Cys GGA Gly	CCTG ATC Ile AGC Ser CCA Pro CAA Gln	AATA CTA Leu GAT Asp AAC Asn CCA Pro	IAAG AGA Arg CTA Leu TTT Phe TGG Trp	TAT Tyr AAA Lys GTC Val TCA Ser	AATC CTG Leu GTG Val ACA Thr AAT ASN	TAAA Lys TAC Tyr GGA Gly TAT Tyr	GGA Gly TGC Cys TTT Phe ACG Thr	GCGA TTA Leu CCA Pro GAT Asp ACC Thr	ICGA Gln GCG Ala ATA Ile TGC	AATT AGG Arg ACA Thr AAT ASN ATT Ile	TTTT GAA Glu TTC Phe AGG Arg GAC Asp	AAAA TGT Cys GAC Asp TTT Phe ATG Met	AAAG GAC Asp GGC Gly GCT Ala GAC Asp	GAAC TTA Leu TGG Trp TTT Phe GAT Asp	114 204 30 294 60 384 90 474 120
В	5' ATG MET AGG Arg TCA Ser CGC Arg CTT Leu	- AC GQA Gly AAG Lys TGT CYS AAA Lys AAG Lys	CACC AAT Asn AGA Arg TGG Trp TGT Cys TTT Phe	TCCG GTA Val GCT Ala <u>AAT</u> ASN TTA Leu CGA Arg	AGTA GAC Asp CAA Gln ACT Thr GAA Glu AAA Lys	CACT AAT Asn TTT Phe ACG Thr AAT ASN GCA Ala	AGTG TTA Leu CAA Gln CCA Pro GGT Gly GTC Val	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ASN	AACA Thr TTA Leu GGT Gly TGG Trp ACT Thr	AAAC. Gln TCG Ser GAA Glu TTT Phe ATA Ile	AGAA Ser ACC Thr ATC Ile CGA Arg TAC Tyr	ATTT. AAT ASN ATT Ile GCT Ala CAT His GTG Val	ACGG CTA Leu TTG Leu TTG Leu CCA Pro GTT Val	CTCC His CCG Pro GCA Ala GAC Asp GGA GLV	CTTC Ser AAT ASN CCG Pro ACT Thr TAT Tyr	IGTAC AAC Asn GTA Val TGT Cys GGA GIY TAC Tyr	CCTG ATC Ile AGC Ser CCA Pro CAA Gln ATA Ile	AATA CTA Leu GAT Asp AAC Asn CCA Pro TCA Ser	TAAG AGA Arg CTA Leu TTT Phe TGG TTP TTT Phe	CACC TAT Tyr AAA Lys GTC Val TCA Ser GCA Ala	AATC CTG Leu GTG Val ACA Thr AAT ASN GCA Ala	TAAA Lys TAC Tyr GGA Gly TAT Tyr TTG Leu	GGA Gly TGC Cys TTT Phe ACG Thr GTC Val	GCGA TTA Leu CCA Pro GAT ASP ACC Thr CTG Leu	ICGA Gln GCG Ala ATA Ile TGC Cys TCA Ser	AATT AGG Arg ACA Thr AAT ASN ATT Ile TTG Leu	TTTT GAA Glu TTC Phe Arg GAC Asp ATA Ile	AAAA TGT GAC Asp TTT Phe ATG Met ATC Ile	AAAGG GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe	GAAC TTA Leu TGG Trp TTT Phe GAT Asp TTG Leu	114 204 30 294 60 384 90 474 120 564 150
В	5' ATG MET AGG Arg TCA Ser CGC Arg CTT Leu ATG <u>Met</u>	- AC GGA Gly AAG Lys TGT VS AAA Lys AAG Lys TTC Phe	CACC AAT ASN AGA Arg TGG Trp TGT CYS TTT Phe AGA Arg	TCCG GTA Val GCT Ala AAT ASN TTA Leu CGA Arg AGC Ser	AGTA GAC Asp CAA Gln ACT Thr GAA Glu AAA Lys CTG Leu	AAT Asn TTT Phe ACG Thr AAT ASN GCA Ala CGA Arg	AGTG TTA Leu CAA Gln CCA Pro GGT Gly GTC Val TGC Cys	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ASN ACT Thr	AACA Thr TTA Leu GGT Gly TGG Trp ACT Thr AGA Arg	AAAC. CAA Gln TCG Ser GAA Glu TTT Phe ATA Ile ATT <u>Ile</u>	AGAA Ser ACC Thr ATC Ile CGA Arg TAC <u>Tyr</u> GCC Ala	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG Val ATC Ile	ACGG CTA Leu TTG Leu TTG CCA Pro GTT Val CAT His	CTCC His CCG Pro GCA Ala GAC Asp GGA GIV GTA Val	CTTC Ser AAT Asn CCG Pro ACT Thr TAT Tyr CAG Gln	IGTAC AAC Asn GTA Val TGT Cys GGA Gly TAC Tyr CTG Leu	CCTG ATC Ile AGC Ser CCA Pro CCA Gln ATA Ile TTT Phe	AATA CTA Leu GAT Asp AAC Asn CCA Pro TCA Ser TCC Ser	TAAG AGA Arg CTA Leu TTT Phe TGG TTP TTT <u>Phe</u> TCG Ser	CACC TAT Tyr AAA Lys GTC Val TCA Ser GCA Ala TTT Phe	AATC CTG Leu GTG Val ACA Thr AAT AAT ASN GCA Ala GCC Ala	TAAA Lys TAC Tyr GGA Gly TAT Tyr TTG Leu GCC Ala	GGA Gly TGC VS TTT Phe ACG Thr GTC Val AAT Asn	GCGA TTA Leu CCA Pro GAT Asp ACC Thr CTG Leu AAT Asn	ICGA Gln GCG Ala ATA Ile TGC CYS TCA Ser CTG Leu	AATT AGG Arg ACA Thr AAT ASN ATT Ile TTG Leu ATG Met	TTTT GAA Glu TTC Phe AGG Arg GAC Asp ATA Ile TGG Trp	AAAA TGT GAC Asp TTT Phe ATG Met ATC Ile ATC Ile	AAAGO GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe ATC Ile	GAAC TTA Leu TGG Trp TTT Phe GAT Asp TTG <u>Leu</u> TGG <u>Trp</u>	114 204 30 294 60 384 90 474 120 564 150 654 180
В	5' ATG MET AGG Arg TCA Ser CTC Leu ATG <u>Met</u> TAT Tyr	- AC GQA Gly AAG Lys TGT AAA Lys TTC AAA Lys TTC Phe AAA Lys	CACC' AAT Asn AGA Arg TGG TTP TGT TGT AGA ACA Arg ACA Thr	TCCG GTA Val GCT Ala AAT Asn TTA Leu CGA Arg AGC Ser GTG GTG Val	AGTA GAC Asp CAA Gln A <u>ACT</u> Thr GAA Glu AAA Lys CTG Leu GTC Val	CACTI AAT Asn TTT Phe ACG Thr AAT AAT AAT AIa CGA AIa CGA AIg GGG Gly	AGTG, TTA Leu CAA Gln CCA Pro GCT GIY GTC Val TGC Cys AAC Asn	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ASN ACT Thr Thr	AACA ACA Thr TTA Leu GGT Gly TGG Trp ACT Thr AGA Arg <u>TCC Ser</u>	AAAC. CAA Gln TCG Ser GAA Glu TTT Phe ATA Ile ATT <u>Ile</u> GTT Val	AGAA TCA Ser ACC Thr ATC Ile CGA Arg TAC <u>Tyr</u> GCC Ala GTA Val	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG Val ATC Ile CAA Gln	ACGG CTA Leu TTG Leu TTG Leu CCA Pro GTT Val CAT His GAA Glu	CTCC CAC His CCG Pro GCA Ala GAC Asp GGA GIV GTA Val AAT Asn	CTTC TCC Ser AAT Asn CCG Pro ACT Thr TAT Tyr CAG Gln CAG	IGTA AAC Asn GTA Val TGT GGA GGA GGA GLY TAC TVT CTG Leu Leu	CCTGJ ATC Ile AGC Ser CCA Pro CCA Gln ATA Ile TTT Phe ATT Ile	AATA' CTA Leu GAT Asp AAC Asn CCA Ser TCC Ser TCC Cys	IAAG AGA Arg CTA Leu TTT Phe TGG TTP TCG Ser CAA Gln	CACC: TAT Tyr AAA Lys GTC Val TCA Ser GCA Ala TTT Phe GTA Val	AATC CTG Leu GTG Val ACA Thr AAT AAT AAT ASN GCA Ala CTG Leu	IAAA AAA Lys TAC Tyr GGA Gly TAT TTG Leu GCC ALA CAT His	CTCCC GGA Gly TGC Val ACG Thr GTC Val AAT AAT ASN GTA Val	GCGA' TTA Leu CCA Pro GAT Asp ACC Thr CTG Leu AAT AAT ATA Ile	CGA Gln GCG Ala ATA Ile TGC CTG CTG Leu TTG Leu	AATT AGG Arg ACA Thr AAT ASN ATT Ile TTG Leu ATG Met CAA Gln	TTTT GAA Glu TTC Phe Arg GAC Arg GAC Arg TIC TTC TTC TTC TTC TTTT	AAAAA TGT GAC Asp TTT Phe ATG ATG Ile ATC Ile TTC Phe	AAAGO GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe ATC Ile ATC Met	GAAC TTA Leu TGG Trp TTT Phe GAT Asp TTG Leu TGG Trp GTG Val	114 204 30 294 60 384 90 474 120 564 150 654 180 744 210
В	5' ATG MET AGG Arg TCA Ser CTC Leu ATG Met TAT Tyr GCC <u>Ala</u>	- AC Gly AAG Lys TGT AAA Lys TTC Phe AAA Lys AAG Lys AAA AAA AAA AAA	CACC ² AAT Asn AGA Arg TGG TTp TGT YS TTT Phe ACA Arg ACA Thr TAC Tyr	TCCG GTA Val GCT Ala AAT TTA Leu CGA Arg CGA AGC Ser GTG Val TTA Leu	AGTA GAC Asp CAA Gln ACT Thr GAA Glu AAA Lys CTG Leu GTC Val TGG Trp	CACT, AAT Asn TTT Phe ACG Thr AAT Asn GCA Ala CGA Ala CGA GIY GGG GIY ATG Met	AGTG, TTA Leu CAA Gln CCA Gly GTC GIY GTC Val TGC Cys AAC ASN TTC Phe	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ASN ACT Thr ACT Thr TGC Cys	AACA ACA Thr TTA Leu GGT GGY TGG Trp ACT Thr AGA Arg <u>TCC</u> Ser GAA Glu	AAAAC. CAA Gln TCG Ser GAA Glu TTT Phe ATA Ile ATT <u>Ile</u> GTT Val GGA GLV	AGAA TCA Ser ACC Thr ATC Ile CGA Arg TAC Tyr GCC Ala GTA Val TTG Leu	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG Val ATC Ile CAA Gln CAT His	ACGG CTA Leu TTG Leu TTG Leu CCA TTG CAT His GAA Glu CTA Leu	CTCC CAC His CCG Pro GCA Ala GAC Asp GGA GIV GTA Val AAT Asn CAT His	CTTC' TCC Ser AAT Asn CCG Pro ACT Thr TAT TVr CAG Gln CTT Leu	IGTAC AAC Asn GTA Val TGT GGA GIY TAC TVT CTG Leu CTG Leu GCA Ala	CCTG ATC Ile AGC Ser CCA Gln ATA Ile TTT Phe ATT Ile CTA Leu	AATA' CTA Leu GAT Asp AAC Asn CCA Pro TCA Ser TCC Ser TCC Cys GTA Val	TAAG AGA Arg CTA Leu TTT Phe TGG TTP TCG Ser CAA Gln GTT Val	CACC; TAT Tyr AAA Lys GTC Val TCA Ser GCA Ala TTT Phe GTA Val GTA Val	AATC: CTG Leu GTG Val ACA Thr AAT ASN GCA Ala GCC Ala CTG Leu TTT Phe	TAAAA AAA Lys TAC Tyr GGA Gly TAT TYr TTG Leu GCC Ala CAT His GTC Val	GTCCC GGA Gly TGC TTT Phe ACG Thr GTC Val AAT ASn GTA Val AAA Lys	GCGA' TTA Leu CCA Pro GAT Asp ACC Thr CTG Leu AAT Asn ATA Ile GAT Asp	CGA CAA Gln GCG Ala ATA Ile TGC CTG Leu TTG Leu GAT Asp	AATT AGG Arg ACA Thr AAT ATT Ile TTG Leu ATG Met CAA Gln TCC Ser	TTTT GAA Glu TTC Phe Arg GAC Asp ATA Ile TGG Trp TAC Tyr GCC Ala	AAAAA TGT GAC Asp TTT Phe ATG ATC Ile ATC Lle ATC Phe ATG Met	AAAAG GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe ATC Ile ATC Met AGG Arg	GAAC TTA Leu TGG Trp TTT Phe GAT Asp TTG Leu TGG Trp GTG Val TGG Trp	114 204 30 294 60 384 90 474 120 564 150 654 180 744 210 834 240
В	5' ATG MET AGG Arg TCA Ser CTT Leu ATG Met TAT Tyr GCC Ala TTC Phe	- AC GQA Gly AAG Lys TGT AAA Lys AAA Lys TTC Phe AAA Lys AAC ASN TAT Tyr	CACC ² AAT Asn AGA Arg TGG Trp TGT TGT Phe ACA Arg ACA Thr TAC Tvr TGT Cvs	TCCG GTA Val GCT Ala AAT Asn TTA Leu CGA Arg AGC Ser GTG Val TTA Leu ATT Ile	AGTA GAC Asp CAA Gln ACT Thr GAA Glu AAA Lys CTG Leu GTC Val TGG Trp GGC Glv	CACT, AAT Asn TTT Phe ACG Thr AAT ASN GCA AIa CGA AIa CGA AIG GIY ATG Met TGG Trp	AGTG, TTA Leu CAA Gln CCA Pro GIY GIY GIY GIY GIY Cys AAC ASN TTC Phe TTT Phe	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ASn ACT Thr ACT Thr TGC CVS CTG Leu	AACA ACA Thr TTA Leu GGT Gly TGG Trp ACT Thr AGA Arg <u>TCC</u> Ser GAA Glu CCG Pro	AAAC. CAA Gln TCG Ser GAA Glu TTT Phe ATA Ile GTT Val GCA GLV GCC Ala	AGAA TCA Ser ACC Thr ATC CGA Arg TAC Tyr GCC Ala GTA GTA TTG Leu ATT ILe	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG Val ATC Ile CAA Gln CAT His TTA Leu	ACGG CTA Leu TTG Leu TTG Leu CCA Pro GTT Val CAT His GAA Glu CTA Leu ACT Thr	CTCC CAC His CCG Pro GCA Ala GAC Asp GGA GIV GTA Val AAT ASN CAT His GCA Ala	CTTC' Ser AAT Asn CCG Pro ACT Thr TAT Tyr CAG Gln CAG Gln CTT Leu ATC Ile	IGTAC AAC Asn GTA Val TGT GGA Gly TAC TYI CTG Leu ITA Leu GCA Ala TAT Tyr	CCTG ATC Ile AGC Ser CCA Pro CAA Gln ATA Ile TTT Phe ATT Ile CTA Leu GCT Ala	AATA CTA Leu GAT Asp AAC Asn CCA Pro TCC Ser TCC Ser TCC Cys GTA Val TGG Trp	TAAG AGA Arg CTA Leu TTT Phe TGG Trp TTT TCG Ser CAA GIN GTT Val GTG Val	CACC: TAT Tyr AAA Lys GTC Val TCA Ser GCA Ala GTA Val GTA Val GTA Val AGA Arg	AATC: CTG Leu GTG Val ACA Thr AAT ASN GCA Ala CTG Leu TTT Phe TCT Ser	IAAA AAA Lys TAC Tyr GGA Gly TAT Tyr TTG Leu GCC Ala GCC Ala GCC Ala	CTCCC GGA Gly TGC TTT Phe ACG Thr GTC Val AAT ASN GTA Lys AAT ASN	GCGA' TTA Leu CCA Pro GAT Asp ACC Thr CTG Leu AAT ASN ATA Ile GAT Asp CCT Pro	CGAA Gln GCG Ala ATA Ile TGC CTG Leu TTG Leu GAT Asp GAT Asp	AATT AGG Arg ACA Thr AAT ASN ATT Ile TTG Leu ATG Met CAA Gln TCC SCI GAT ASP	TTTT GAA Glu TTC Phe Arg GAC Asp ATA Ile TGG Trp TAC Tyr GCC Ala ACA Thr	AAAAA TGT GAC Asp TTT Phe ATG Met ATC Ile TTC Phe ATG Met ATG Met AGG Arg	AAAAG GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe ATC Ile ATC ATG Met AGG Arg CIn	SAAC TTA Leu TGG Trp TTT Phe GAT Asp TTG Leu TGG TTD GTG GTG Val TGG TTD TGT Cys	114 204 30 294 60 384 90 474 120 564 150 654 180 744 210 834 240 924 270
В	5' ATG MET AGG Arg CCA Ser CGC Arg CTT Leu ATG Met TAT Tyr GCC Ala TTC Phe TGG Trp	- AC GGA Gly AAG Lys TGT AAA Lys AAA Lys AAA Lys TTC Phe AAA Lys AAC ASN TAT Tyr ATG Met	CACC' AAT Asn AGA Arg TGG TTP TGT TGT ACA Arg ACA Arg ACA Asn	TCCG GTA Val GCT Ala AAT TTA Leu CGA Arg AGC Ser GTG Val TTA Leu ATT Ile GAA Glu	AGTA GAC Asp CAA Gln ACT Thr GAA Glu AAA Lys CTG GIC Val TGG Trp GGC Gly TCG Ser	AACT AASN TTT Phe ACG Thr AAT AAT AAT AAT CGA AIa CGA AIa CGA AIa CGA ATG GIY ATG TTG TTP TAT TYr	AGTG. TTA Leu CAA Gln CCA Pro GGT GIY GTC Val TGC Cys AAC Asn TTC Phe ACA Thr	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ASN ACT Thr ACC Thr TGC CVS CTG Leu CAA Gln	AACA ACA Thr TTA Leu GGT Gly TGG Trp ACT Thr AGA Arg TCC Ser GAA Glu CCG Pro TGG Trp	AAAAC. CAA Gln TCG Ser GAA Glu TTT Phe ATA Ile ATA Ile GTT Val GGA Gly GCC Ala ATT Ile	AGAA Ser ACC Thr ATC Ile CGA Arg TAC Tyr GCC Ala GTA Val TTG Leu ATT Ile CTG Leu	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG Val ATC Ile CAA Gln CAT His TTA Leu ATT Ile	ACGG CTA Leu TTG Leu TTG Leu CCA Pro GTT Val CAT His GAA Glu CTA Leu ACT Thr GTA Val	CTCC CAC His CCG Pro GCA Ala GAC Asp GGA GIV GTA Val AAT Asn CAT His GCA Ala CCA Pro	CTTCC Ser AAT Asn CCG Pro ACT Thr TAT TYr CAG Gln CAG Gln CTT Leu ATC Ile GTT Val	rGTAG AAC Asn GTA Val TGT GGA GGA GGA GGA TAC TYT CTG Leu GCA Ala TAT TYT TGT Cys	CCTGI ATC Ile AGC Ser CCA Pro CCA Gln ATA Ile TTT Phe ATT Ile CTA Leu GCT Ala TTG Leu	AATA' CTA GAT Asp AAC Asn CCA Pro TCA Ser TCC Ser TCC Cys GTA TCC GTA TGG TTp TCA Ser	TAAGA AGA Arg CTA Leu TTT Phe TGG TTP TCG Ser CAA Gln GTT Val GTG Val TTA Leu	CACCJ TAT Tyr AAA Lys GTC Val TCA Ser GCA Ala GTA Val GTA Val GTA Val AGA Arg TTT Phe	AATC CTG Leu GTG Val ACA Thr AAT AAT AAT AAT AAT CTG Leu TTT Phe TCT Ser GCT Ala	IAAAA AAA Lys TAC Tyr GGA Gly TAT TTG Leu GCC Ala GCC Ala GCC Ala GCC Ala Ala AGC Ser	CTCCC GGA Gly TGC CYA TTT Phe ACG Thr GTC Val AAT AAT AAA Lys AAT ASN CTT Leu	GCGA! TTA Leu CCA Pro GAT Asp ACC Thr CTG Leu AAT ASP AAT ASP CCT Pro GGA GLV	CGAA Gln GCG Ala ATA Ile TGC CTG Leu TTG CTG Leu GAT Asp GAT Asp TTT Phe	AATT AGG Arg ACA Thr AAT AAT ATT Ile TTG Leu ATG Met CAA Gln TCC Sef GAT Asp CTA Leu	TTTT GAA Glu TTC Phe AGG Arg GAC Asp ATA Ile TGG Trp TAC Tyr GCC Ala ACA Thr ATC Ile	AAAAA TGT GAC Asp TTT Phe ATG Met ATC Ile TTC Phe ATC Ile ATC ATC ATC AATG AATG AAT AAT	AAAGG GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe ATC Ile ATG ATG ATG GIn GTG Val	SAAC TTA Leu TGG Trp TTT Phe GAT Asp TTG Leu TGG TTG GTG GTG Cys GTT Cys GTT Val	1114 204 30 384 90 474 120 5564 180 744 210 834 240 924 270 924 270
В	5' ATG MET AGG Arg TCA Ser CLC CAC ATG Met TAT Tyr GCC Ala TCC Phe TGG Trp CGA	- AC GQL AAG Lys TGT AAA Lys AAA Lys TTC Phe AAA Lys AAC ASN TAT Tyr AAG Met GTA Val	CACC' AAT AASA AGA AGA TGG TTP TGT TGT ACA ATA TAC ACA ACA ASA CVS AAC ASA	TCCG GTA Val GCT Ala AAT TTA Leu CGA Arg AGC Ser GTG Val TTA Leu ATT Ile GAA Glu CTA Leu	AGTA GAC Asp CAA Gln ACT Thr GAA Glu AAA Lys CTG CTG CTG CTG CTG GIV TCG Ser ACG Thr	CACT: AAT Asn TTT Phe ACG Thr AAT AAT AAT AAT CGA ATG GIY AATG TTP TAT Tyr AAA Lys	AGTG. TTA Leu CAA Gln CCA Pro GGT GTC Val TGC Cys AAC Asn TTC Phe ACA Thr TTA Leu	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ACT Thr ACC Thr TGC Cys CTG Leu CAA Gln CAT His	AACA ACA Thr TTA Leu GGT Gly TGG Trp ACT TCC Ser GAA Glu CCG Pro TGG Trp TGG Trp	AAAAC. CAA Gln TCG GAA Glu TTT Phe ATA Ile GTT Val GGA GIV GGA ATT Ile AATT LIE	AGAA TCA Ser ACC Thr ATC CGA Arg TAC Tyr GCC Ala GTA Val TTG Leu ATT Ile CTG Leu TCT Ser	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG Val ATC Ile CAA Gln CAT His TTA Leu ATT Ile GCC Ala	ACGG CTA Leu TTG Leu CCA Pro GTT Val CAT His GAA Glu CTA Leu ACT Thr GTA Val AAT Asn	CTCC CAC His CCG Pro GCA Ala GAC ASP GGA GIV Val AAT ASN CAT His GCA Ala CAT Pro CCT Pro	CTTCC Ser AnT Asn CCG Pro ACT Thr TAT Tyr CAG Gln CTT Leu ATC Lle GTT Val GCA Ala	rGTAG AAC AASn GTA Val TGT GGA Gly TAC TYr CTG Leu CTG Leu CTG Leu CTG Leu CTG CA Ala TAT Tyr TGT Cys CCG Pro	CCTG/ ATC Ile AGC Ser CCA CTA CAA GIN ATA Ile CTA Leu GCT Ala TTG Leu GTT Val	AATA' CTA GAT GAT Asp AAC Asp CCA Pro TCA Ser TCC Cys GTA TCC Cys GTA TCC Cys GTA TCC Cys GTA TCC Cys GTA	TAAG AGA Arg CTA Leu TTT Phe TCG Ser CAA GIN GTT Val GTG Val TTA Leu CTT Leu	CACC. TAT Tyr AAA Lys GTC Val TCA Ser GCA Ala TTT Phe GTA Val GTA Val GTA Val AGA Arg TTT Phe AGG Arg	AATC' CTG Leu GTG Val ACA Thr AAT AAT AAT CTG Leu TTT Phe TCT Ser GCT Ala AAG Lvs	IAAAA AAA Lys TAC Tyr GGA Gly TAT TTG Leu GCC Ala GCC Ala AGC Ser GCT Ala	CTCCC GGA TGC TGC TTT Phe ACG Thr AAG Thr AAT AAT AAA Lys AAT AAA Lys AAT CTT Leu GTA Val	GCGA TTA Leu CCA Pro GAT Asp Acc Thr CTG Leu AAT ATA Ile GAT Asp CCT Pro GGA GIV CGA Arq	CGAA Gln GCG Ala ATA Ile TGC CTG Leu TTG CTG Leu GAT Asp GAT Asp TTT Phe GCG Ala	AATT AGG Arg ACA Thr AAT AAT AAT TTG Leu ATG ATG GAT CCA GAT ASp CTA Leu GCG Ala	TTTTT GAA Glu TTC Phe GAC GAC Asp ATA Ile TTB TGG TTB TAC TYr GCC Ala ACA Thr ATC Ile CTA Leu	AAAAA TGT GAC Asp TTT Phe ATG Met ATC Ile Ilc ATC Phe ATC ATC ATC AATG AATG AAT AATA Ile	AAAGG GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe ATC Ile ATG ATG Gln GIn GTG Val CTG Leu	SAAC TTA Leu TGG TTP Phe AASp TTG Leu TGG TTG GTG Cys GTG Cys GTT TGT Cys GTT TGT Cys GTA Val	114 204 30 294 60 474 150 564 150 654 180 744 210 834 210 834 210 1014 300
В	5' ATG MET AGG Arg CCA Ser CCA Arg CTT Leu ATG Met TAT Tyr GCC Alta TTC Phe TGG Trp CGA Arg	- AC Galy AAG Lys TGT AAA Lys AAA Lys TTC Phe AAA Lys AAC ASN TAT Tyr ATG Met GTA Val TTG Leu	CACC' AAT Asn AGA Arg TGG Trp TGT TGT ACA Arg ACA Arg ACA Thr TAC Cys AAC Asn CLeu TTT Phe	TCCG GTA Val GCT Ala AAT Asn TTA Leu CGA Arg AGC Ser GTG GTG TTA Leu TTA Leu CTA Leu CGA ATT CLeu GAA Glu CTA CTA	AGTAI GAC Asp CAA Gln Thr GAA Glu AAA Lys CTG GLU TGG GC Val TGG Glv TCG Glv TCG Ser ACG Thr ACT TIG	CACTI AAT Asn TTT Phe ACG Thr AAT Asn GCA Ala CGA Ala CGA Arg GGG Gly CGA Trg TAT TGG TAT Tyr AAA Lys CAT His	AGTG. TTA Leu CAA Gln CCA Pro GGT Gly GTC Val TGC Cys AAC ASN TTC Phe ACA Thr Tha Leu CAA ACA ASN TTC Cys ACA ACA ACA ACA ACA ACA ACA AC	ACGC. ACG Thr Tyr TCC Ser ACT Thr AAC ASN ACT Thr AAC ASN ACT Thr CCC CVC CVC CVC CLeu CAA Gln CAT His ATA	AACA ACA Thr TTA Leu GGT Gly TGG Trp ACT Thr AGA Arg TCC Ser GAA Glu CCG GAA Glu CCG Trp TGG Trp TGG Trp TGG Cys CCS CCG Leu	AAAAC. CAA GIn TCG Ser GAA Glu TTT Phe ATA Ile ATT Ile GGT Val GGA GIV GGA ATT Ile ATT ILE AAT AAT AAT ATA	AGAA TCA Ser ACC Thr ATC CGA Arg TAC TVr GAC Arg Arg Arg Arg Arg Arg CGA GCA CA CTG Leu TCT Ser CCA Pro	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG Val ATC Ile CAA Gln CAT HIS CAA CAT HIS CAA ATT ILE GCC ALT ATT Phe	ACGG CTA Leu TTG Leu CCA Pro GTT Val CAT His GAA Glu CTA Leu CTA Leu CTA Leu CTA Leu CTA AAT AAT AAT AAT	CTCCC CAC His CCG Pro GCA Ala GAC Asp GGA GIV Val AAT Asn CAT His CAT His CAT Pro CCT Pro CCT Pro CCA Pro	TTCC Ser AAT Asn CCG Pro ACT Thr TAT Tyr CAG Gln CTT Leu CAG Gln CTT Leu CAG Gln CTT Leu GTT Val ACA ALA GAG Glu	rGTAG AAC Asn GTA Val TGT GGA Gly TAC TYT CTG Leu ITA Leu GCA Ala TAT TYT CYS CCG Pro CCG Pro	CCTGJ ATC Ile AGC Ser CCA Fro CCA Gln ATA Ile TTT Phe ATT Ile CTA Leu GCT TTG Leu GTT Val AAA Lys	AATA' CTA Leu GAT Asp AAC Asn CCA Pro TCA Ser TCC Ser TCC Ser TGC Cys GTA TGG TGG Trp TCA Ser GIY GCA Ala	TAAGA AGA Arg CTA Leu TTT Phe TGG Trp TTT Phe CAA GIN GTT CAA GIN GTT CAA GTG GTG GTG GTG CAA CTT Leu CCT Leu	CACC: TAT Tyr AAA Lys GTC Val TCA Ser GCA Ala GTA Val GTA Val GTA Val GTA Val GTA Val GTA AGG Arg GGA AGG GIY	AATC: CTG Leu GTG Val ACA Thr AAT ASN GCA Ala CTG Leu TTT Phe TCT Ser GCT Ala AGC ALa UTT TCT Ser ALa AAG Luys GAA GIU	TAAAA ALAA Lys Tyr GGA Gly TAT Tyr TTG Leu GCC Ala GCC Ala AGC Ser Ala AGC CAA Ala AGC CAA	CTCCC GGA Gly TGC TTT Phe ACG Thr GTC Val AAT AAT AAT AAA Lys CTT Leu CTT Leu GTA Val	GCGA' TLAU CCA Pro GAT Asp ACC Thr CTG Leu AAT ASP ATA ATA SP CCT GGA GGA GGA CCT CTO GGA CCGA Arq TAT	CGAA GAA GCG Ala ATA Ile TGC CTG CTG CTG CTG Leu TTG GAT ASp TTT Phe GAT ASp GAT ASp GAT GAA GCG GAA GIU	AATT AGG Arg ACA Thr AAT ASn ATT Ile TTG Leu ATG GIN TCC CAA GIN TCC GAT ASp CTA Leu GCG GAI AIA	GAA Glu TTC Phe AG GAC Asp ATA Ile TGG Trp TAC Tyr GCC Ala ACA Thr ATC Ile CTA Leu TTT Phe	AAAAA TGT GGC Asp TTTT Phe ATG Met ATG Ile ATC Ile ATC Ile ATC AGG Arg AAT AATA AILE TCA Ser	AAAGC GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe ATC Ile ATC GC Gln GTG Val CTG Leu GCA Ala	SAAC TTA Leu TGG Trp Phe GAT Asp TTG GAT Asp TTG GTG TTC GTG Cys GTT Cys GTT Cys GTT QTA Cys Cys Cys Cys Cys Cys Cys CTT Leu CTT Leu	114 204 30 294 60 474 120 564 150 654 150 654 210 834 240 924 270 1014 300 1104 330
В	5' AGG ATG ATG ATG CTA Ser CGC ATG CTT Leu ATG ATG TTT TTT TTT CGA ATG CCA TTC CGA ATG CCA CCA CCA CCA CCA CTT CGC CCA CTT CCA CCA CTT CA CCA CTT CA CCA C	- AC GA GLy AAG Lys TGT Lys AAA Lys TTC Phe AAA Lys AAC ASA AAC ASA AAC ASA TAT TVT TVT GTA TG GTA TG GTT TCQ GTT	CACC' AAT Asn AGA Arg TGG Trp TGT TGT ACA TTP Phe AGA Arg AcA Arg AcA Thr TGT Cys ACA Asn CTC Leu TTP Phe	TCCG GTA Val GCT Ala AAT TTA Leu CGA Arg CGA Arg Val TTA Leu ATT Ile GAA Glu CTA Leu GGT GTY Leu	AGTAI GAC Asp CAA Gln ACT Thr GAA Glu AAA Lys CTG Ucu CTG CTC Ucu CTG G1V TCG Ser ACG Chr ACG Thr ACG	CACTJ AAT AASN TTTT Phe ACG Thr AAT AAT AAT AAT AAT AAT AAT AAT CGA ATG GGG GGG GGG GGG CA TTG TTP TAT TTP TAT TYT AAA AAS AAT AAT AAT AAS AAT AAT	AGTG, TTA Leu CAA Gln CCA Pro GGT GIY GTC Val TGC Cys Aac TTC Phe TTT Phe ACA Thr TTA Leu CACA TTTA Leu	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ASR ACT Thr AAC CVS CTG Leu CAA Gln CAT His ATA Thr ACT Thr TGC CVS	AACA AACA Thr TTA Leu GGT TGG TTP ACT Thr AGA Arg CCT GAA Glu CCG Pro TGG TTP TGG TTP TGC CYS CTG CYS CTG Leu Val	AAAAC GAA GIN TCG Ser GAA Glu TTT Phe ATA Ile ATA Ile GGA GIV GGA AIA ATT Ile AAT AAT AAT AAT ATA	AGAA TCA Ser ACC Thr ATC Ile CGA Arg TAC Tyr GCC Ala GTA Val TTG Leu TTG Leu TCT Ser CCA Pro GTA Val	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG CAT His CAA Gln CAT His TTA Leu ATT Ile GCC CAA ATC CAT His CAA CAT His CAA CAT His CAT CAT CAT CAT CAT CAT CAT CAT CAT CAT	ACGG CTA Leu TTG Leu TTG Leu CCA TTG CAT His GAA Glu CTA Leu ACT Thr GTA Val AAT AST CAG AAT AST CAG AAT TTT Phe	CTCCC CAC His CCG Pro GCA Ala GAC Asp GGA GIA Val AAT Asn CAT His GCA Ala CAT His CCT Pro CCT Pro CCT Pro CCA CCA CCA CCA CCA CCA CCA CCA CCA CC	TTCC Ser AAT Asn CCG Pro ACT Thr TAT TAT CAG GIn CTT Leu ATC GIN CTT Leu ATC Ile GTT Val GCA GAG GAG GAG GIU TTT Phe	rGTA(AAC Asn GTA Val TGT GGA Gly TAC CLGU TAC Leu ITA Leu GCA Ala TAT TYT TGT CVS CCG Pro CCG Pro GCC Val	CCTGJ ATC Ile Ser CCA Ser CCA Gln ATA Ile ATT Ile CTA Leu GCT ALA GTT TTG Leu GCT ALA ALA ALA ALA AAA	AATA: CTA Leu GAT Asp AAC Asn CCA Ser TCA TCA TCC STC Cys GTA TGC Cys GTA TGG GTA TCA Ser TCA Ser CA Ser CA SCA TCA CA SCA TCA CA SCA CA CA SCA S	TAAGA Arg CTA Leu TTT Phe TGG TTP TTT Phe TCG CAA GIN GTT Val GTG Val GTG Val CTT Leu CTT Leu CTT Leu	CACCI TAT Tyr AAA Lys GTC Val TCA Ser GCA Alt GTA Val GTA Val GTA AGG Arg GGA Arg GGA GIY GTG Val	AATC: CTG Leu GTG Val ACA Thr AAT GCA Ala GCC Ala CTG Leu TTT Phe Ser GCT Ala AAG Lys GAA GLUS	TAAAA Lys TAC Tyr GGA Gly TAT TTG Leu GCA Ala GCC CAT His GCC CAT Ala AGC Ser GCT ALa CGA ALa CGA ALA CGA ALA CGA ALA CGA CGA CCA CCA CCA CCA CCA CCA CCA CC	CTCCC GGA Gly TGC TTT Phe ACG Thr GTC Val AAT Asn Val AAA Lys AAT Asn CTT Leu GTA Val GTA Val GTA AAA	GCGA' TTA Leu CCA Pro GAT Asp ACC Thr CTG Leu AAT ASp CCT Pro GGA GIV CGA Ard TAT Tyr Tyr TCC Phe	CGAA Gln GCG Ala ATA Ile TGC CTG CTG CTG CTG CTG CTG CTG CTG CTG	AATT AGG ACA Thr AAT ASN ATT Lle TTG Leu TTG CAA GCA GAT ASP CTA Leu GCG Ala	GAA GAU TTC Phe GAC Arg GAC Asp ATA TTG TGG TTG TTG TTG TTG TAC TTG Ala ACA Thr ATC Leu TTT E ATG ATG ALA ACA	AAAA TGT GGC Asp TTT Phe ATG Met ATC Ile TTC Phe ATG Met ATG ATG ATG ATG ATG ATG ATG ATG CA Ser TCA	AAAAG GAC Asp GGC Gly GCT Ala GAC Phe ATC Phe ATC Ile ATG Met ATG GIN GTG CTG Leu GCA Ala CGA AAg	SAAC TTA Leu TGG Trp Phe GAT Asp TTG GAT TTG TGG TTD TGG TGG TGG TTD TGG TGG CYs GTA Val GTA Val CTT Leu CTT CSS GTA	114 204 30 294 60 474 120 564 150 654 180 744 210 834 240 924 270 1014 300 1104 330 1194
В	5' ATG MET AGG Arg TCA Ser CLA CGC Arg TTC TTC TTC TTC TTC TTC TTC TTC TTC CGA Arg CCA TTC TCG TTC CGA Arg TCC ALL TTC TCC TCC CCC CCC CCC CCC CCC CCC C	- AC Gly AAG Lys TGT AAA Lys AAA Lys TTC Phe ALys AAC ASN TAT TYr AAG GTA Val CAA GTT Val	CACC' AAT Asn AGA Arg TGG TTP Phe ACA Arg ACA Thr TAC CYS AAC Asn CTC Leu TTT Phe CTC CYS CTC Leu TTT Phe CTC CTC ASN	TCCG GTA Val GCT Ala AAT Asn TTA Leu CGA Arg GTG Val TTA Leu CTA Leu CTA Leu CTA Leu CTA Leu CTA Lau CTA ATT TTA Lau CTA ATT CIA	AGTA' GAC Asp CAA Gln ACT Thr GAA Glu ALys CTG Leu GTC Val TCG GGC Val TCG GGC Trp TCG Ser ACG Thr ACG Thr ACG Thr CAG AAA CGC AAA CI CAA Glu CAA CI CI CAA CI CI CAA CI CI CAA CI CI CAA CI CI CI CI CI CI CI CI CI CI CI CI CI	CACT: AAT Asn TTT Phe ACG Thr AAT GCA ATG GCA ATG GGG G1y CGA ATG TGG TGG TGG TGG TGG TAT Tyr AAA Lys CAT His GCA Ala	AGTG, TTA Leu CAA Gln CCA Pro GGT GGT Val TGC Cys AAC Asn TTC Phe GAT Asp	ACGC. ACG Thr TAT Tyr TCC Ser ACT Thr AAC ACT Thr AAC ACT Thr CAS CLGU CAA Gln CAT His ATA Ile TGT Cys AAT AAS	AACA ACA Thr TTA Leu GGT Gly TGG Trp ACT Thr ACA Arg TCC Ser GAA Glu CCG Cys CCG TCG TGG Cys CLG GTA GCT GGT GGT GGT GGT	AAAAC CAA Gln TCG Ser GAA Glu TTT Phe ATA Ile GTT Val GGA GIV GCC AATT Ile AAT ASN ATA ATA ATA ATA	AGAA Ser Thr ACC Thr ATC CGA Arg TAC TVr GCC Ala GTA TTG GCC Ala TTG Leu TCT Ser CCA Pro GTA Val TTG Leu	ATTT. AAT Asn ATT Ile GCT Ala CAT His GTG Val ATC Ile CAA Gln CAT His CAA Gln CAT TTA Leu ATT Ile GCC Ala ATT ILE CAA ATT ILE CAA ATT LEU CAA CAT CAT CAT ASN ATT AST ASN ATT ATT ASN ATT ASN ATT AST ASN ATT ASN ATT ASN ATT ASN AST ASN ASN AST ASN ASN ASN ASN ASN ASN ASN ASN ASN ASN	ACGG CTA Leu TTG Leu TTG CAC CTA CAT His GAA Glu CTA Leu CTA Leu AACT TTT GTA Val AAT ASN CGA AAT TTT TTP GCA Ala	CTCCC CAC His CCG Pro GCA Ala GAC GGA GGA GIV Val AAT ASN CAT His CCA Pro CCA Pro CCA Pro CCA TGT CCA TGT CYS	CTTCC Ser AAT Asn CCG Pro ACT Thr TAT TVT CAG Gln CAG Gln CTT Leu CAG Gln CTT Leu CAG GlT TTC CAG GlT TTC CAG GlT CTT Leu CAG CG CAA Gln CTT CCG CAA CCG CAA CCG CCG CAA CCG CCG CAA CCG CCG	rGTA(AAC Asn GTA Val TGT GGA GGA GGA TAC TYr TA Leu GCA Ala TAT TAT TAT TGT CVS CCG Pro CCG Pro CCG GTC Val ACC Thr	CCTGJ ATC Ile AGC Ser CCA Pro CAA ATA Ile TTT Phe ATT Ile CTA Leu CTA Leu GTT Val AAA Lys CAA AAGA AGA	AATA' CTA Leu GAT Asp AAC Asp AAC Asn TCA Ser TCC Ser TCC Ser TCC GTA TCC GTA TCC GTA TCA Ser TCA Ser TCA Ser GCA Ala GTC GCA Ala GTC GCA	TAAGA AGA Arg CTA Leu TTT Phe TGG Trp TTT Phe CAA GIN GTT CAA GIN GTT Ual CTA Leu CTT Leu CTT Leu CTA CAA	CACCI TAT Tyr AAA Lys GTC Val TCA GCA Ala TTT Phe GTA Val GTA Val GTA Val AGA AGG Arg GGA GIY GTG CA ATG Met	AATC: CTG Leu GTG Val ACA AThr AAT AAT AAT AAT AAT CTG CTG CTG CTG CTG CTG CTG CTG CTG CT	TAAAA AAA Lys TAC Tyr GGA Gly TAT TTT GCC Ala CAT His GCC Ala CAT His GCC CAT ACC CAT ACC CCC ACC CCC ACC CCC C	CTCCC GGA Gly TGC TTT Phe ACG TTT Phe GTC Val AAT AAT AAT AAA CTT Leu GTA Val GTA Val AAA CTT Leu GCT AAT AAA AAA CTT AAA AAA CTT AAA AAA CTT AAA AAA	GCGA' TTA Leu CCA Pro GAT Asp CTG Leu AAT ASD ATA Ile GAT ASD CTG GGA GIV CGA Arq TAT Tyr TTC Phe TAGA	rcGA Gln GCG Ala ATA Ile TGC CTG Leu TTG CTG Leu TTG GAT Asp GAT Asp GAT Asp GAT Ala GAA GIU AAA GIU AAA GIU	AATT AGG Arg ACA Thr AAT ATT Ile TTG Leu ATG GIn TCC GAA GIN TCCC GAA GIN CTA Leu GCG Ala ATA Ile GCG Ala	TTTT GAA Glu TTC Phe Arg GAC Arg Arg Trg TaG Trp TAC Tyr TAC Tyr ATA ACA ATA Leu Thr ATC Leu TTT Phe ATG GCC Ala ACA A ATA Thr GCC ACA ASP TTC TYR TAC TCC ASP ATA TCC TCC ASP ATA TCC TCC ASP ATA TCC TCC ASP ATA TCC TCC ASP ATA TCC TCC ASP ATA TCC TCC ASP ATA TCC TCC ASP ATA TCC TCC ASP ATA TCC TCC ASP ATA TCC TCC TCC ASP ATA TCC TCC TCC ASP ATA TCC TCC TCC TCC TCC TCC ASP ATA TCC TCC TCC TCC TCC TCC TCC TCC TCC	AAAAA TGT GAC Asp TTT Phe ATG ATG ATG ATG ATC Ile TTC Phe ATG AAT AAT AAG AAG AAT AAT AC AGG AAT AC CA Ser GTA Val	AAAAGG GAC Asp GGC Gly GCT Ala GAC Asp TTC Phe ATC Ile ATC Ile ATC GIN GCA ATG GIN GTG CTG Leu CTG Leu CCA Ala CGA Arg ATG GTG CTG CTG CTG CTG CTG CTG CTG CTG C	SAAC TTA Leu TGG Trp Phe GAT Asp TTG Leu TTG GTG TTD TGG TTG TGG TGG TGG TGG TGG	1114 204 30 294 60 384 90 474 120 564 150 654 150 654 120 924 270 1014 300 1104 300 1194 360 1382

Figure 1. cDNA sequences and the deduced amino acid sequences of Rhopr-CT/DH-Rs. *Rhopr-CT/DH-R1-B* (A) and *Rhopr-CT/DH-R2-B* (B). The numbering for each sequence is shown at right. Within the nucleotide sequence, the exon-exon boundaries are shaded in gray and the potential polyadenylation signal is double-underlined. Within the amino acid sequence, the initial methionine start codon has been capitalized, the six conserved cysteine residues are shaded in red, the potential phosphorylation sites are shaded in black, the potential N-linked glycosylation sites are boxed and the seven predicted transmembrane domains are underlined. The two amino acid residues (valine and serine) that are absent in Rhopr-CT/DH-R1-C are dash underlined in Rhopr-CT/DH-R1-B.

doi: 10.1371/journal.pone.0082466.g001



Figure 2. *Rhopr-CT/DH-R1* and *Rhopr-CT/DH-R2* splicing. Molecular organization of *Rhopr-CT/DH-R1* (A) and *Rhopr-CT/DH-R2* (B) splice variants based on BLAST analysis and splice site prediction. The boxes represent exons (drawn to scale). The regions shaded in gray represent the open reading frame while the unshaded regions represent the untranslated regions. The dashed boxes represent predicted regions that were not cloned. doi: 10.1371/journal.pone.0082466.g002

comprising CT/DH-Rs, CT/DH-R1s form a monophyletic group and CT/DH-R2s form another monophyletic group, which suggests that these two receptor subtypes arose from a recent duplication in insects, independent from the one in deuterostomes.

Spatial expression profile of Rhopr-CT/DH-R transcript variants

In order to discover physiological targets of Rhopr-CT/DH, qPCR analysis was performed to determine the spatial expression pattern of Rhopr-CT/DH-R transcript variants (Figure 6). Rhopr-CT/DH-R1-A, despite lacking the characteristic seven transmembrane domains, is highly expressed in the testes (Figure 6A). Rhopr-CT/DH-R1-B/C, on the other hand, is highly enriched in the CNS and dorsal vessel and expressed at lower levels in the foregut, salivary glands, hindgut, testes, ovaries and prothoracic glands. Rhopr-CT/DH-R2-A is only expressed in the CNS and in low amounts (Figure 6B). Rhopr-CT/DH-R2-B has the highest abundance in the CNS and this is over 600 fold higher compared to any other Rhopr-CT/DH-R transcript levels in any tissue. Interestingly, Rhopr-CT/DH-R2-B and not Rhopr-CT/DH-R1-B/C, is expressed in MTs. Rhopr-CT/DH-R2-B is also expressed at much lower levels in the salivary glands, testes, ovaries and prothoracic glands. With regards to the adult reproductive tissues, both *Rhopr-CT/DH-R1-B/C* and *Rhopr-CT/DH-R2-B* have the highest expression in ovaries and are expressed at lower levels in the testes and female reproductive tissues minus the ovaries (Figure 7). Similar to the fifth-instar, *Rhopr-CT/DH-R1-A* is highly expressed in the adult testes (Figure 7A).

Discussion

In the present study, we have found that *R. prolixus*, and perhaps other insects, contain two CT/DH-Rs. *Rhopr-CT/DH-R1* encodes three splice variants while *Rhopr-CT/DH-R2* encodes two splice variants. We functionally characterized full-length receptor isoforms encoded by these transcript variants in HEK293/CNG cells. The assay used to characterize these receptors monitored calcium mobilization into the cells from the extracellular media. Since Rhopr-CT/DH is thought to mediate its effects via cAMP as the secondary messenger, its binding to the receptor would result in receptor activation and a subsequent increase in intracellular cAMP levels [7]. This cAMP would then bind to the CNG channel, resulting in its opening and an influx of calcium from the extracellular medium. The amount of calcium mobilized was detected using the



Figure 3. Functional assay of *R. prolixus* CT/DH receptor isoforms (Rhopr-CT/DH-R1-B, Rhopr-CT/DH-R1-C and Rhopr-CT/DH-R2-B) transiently expressed in HEK293/CNG cell lines.

Dose-dependent effect on the bioluminescence response after addition of Rhopr-CT/DH to HEK293/CNG cells expressing Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C (A) and Rhopr-CT/DH-R2 (B). Vertical bars represent SEM (n=3). Rhopr-CT/DH-R2-B is much more sensitive to Rhopr-CT/DH (EC_{50} = 15nM) compared to Rhopr-CT/DH-R1-B or C (EC_{50} = 200-300nM) and results in a greater response.

doi: 10.1371/journal.pone.0082466.g003

reporter, aequorin. Using this assay, we confirmed that Rhopr-CT/DH-R1-B, Rhopr-CT/DH-R1-C and Rhopr-CT/DH-R2-B were all activated by Rhopr-CT/DH. This work therefore essentially de-orphans the D. melanogaster hector and related receptors in other insects. Responses that were robust (ranging from 42 to 191-fold higher than basal response) and sensitive (EC₅₀ values in the low nanomolar range) were obtained in this cell line. This is surprising because according to a previous report, HEK293 cells expressing Drome-CT/DH-R1 were not responsive to Drome-CT/DH until RCP, an accessory protein, was co-expressed with the receptor [12]. Even after the expression of D. melanogaster RCP along with the receptor, the maximum response obtained was under 2-fold that of the basal response. An improved sensitivity (EC₅₀ = 82nM) and a greater response (7-fold) to Drome-CT/DH was only observed when the human RCP and RAMPs were expressed along with the receptor. This led to the conclusion that perhaps RCP and RAMPs are required for signalling through insect CT/DH-Rs just like mammalian calcitonin receptors: however, our results with a similar cell line, coupled with the fact that proteins with sequence homology to human RCP and RAMPs are not found in the R. prolixus genome, suggests that Rhopr-CT/DH-Rs may not require accessory proteins for effective signalling.

Comparing the responses of the two Rhopr-CT/DH-Rs, Rhopr-CT/DH-R2-B is at least 10-fold more sensitive to Rhopr-CT/DH than are Rhopr-CT/DH-R1-B or Rhopr-CT/DH-R1-C; the EC₅₀ values of Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C range from 150-300nM. These values are still relatively high and hence we don't rule out the possibility that other

endogenous ligands may also activate this receptor. Another possibility is that Rhopr-CT/DH-R1 and Rhopr-CT/DH-R2 may interact as heterodimers as is so common of family B1 GPCRs [30]. Either way, Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C do not seem to be functionally different. Rhopr-CT/DH-R1-A is highly expressed in the testes and such high expression may suggest a role for this truncated receptor isoform rather than it being just an error in splicing. Since Rhopr-CT/DH-R1-A and Rhopr-CT/DH-R1-B/C are expressed in the CNS, dorsal vessel, testes and ovaries, we questioned whether Rhopr-CT/DH-R1-A could interact with Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C and affect their signalling. If Rhopr-CT/DH-R1-A influenced (either inhibited or stimulated) the surface expression of Rhopr-CT/DH-R1-B/C, one would expect their EC₅₀ values to be altered. However, Rhopr-CT/DH-R1-A did not influence the signalling through Rhopr-CT/DH-R1-B and Rhopr-CT/DH-R1-C in our heterologous assay utilizing CHO/G16 cells. Hence, the role of Rhopr-CT/DH-R1-A, if any, is still unclear.

Spatial expression analysis of the two receptors using qPCR demonstrates that they have distinct expression patterns, with both receptors expressed centrally and peripherally. We utilized *Rhopr-CT/DH-R1-B* transcript levels in salivary glands and testes cDNA to show the relative expression of each variant for both the receptors in fifth-instar tissues and adult reproductive tissues, respectively. We acknowledge that this method is not as effective as absolute quantification to compare the expression of two different genes and hence this comparison between the two receptors is only an approximate. Nonetheless, at least one of *Rhopr-CT/DH-R1-B*, *Rhopr-*



Figure 4. Multiple sequence alignment of select insect CT/DH receptors. Identical and similar amino acids across 50% of the sequences have been highlighted in black and gray, respectively. The six highly-conserved cysteine residues in N-terminal domain have been marked with an asterisk. Two N-linked glycosylation sites which are conserved across all sequences have been highlighted in red. The predicted locations of the seven transmembrane domains of Rhopr-CT/DH-R1 and Rhopr-CT/DH-R2 have been indicated using green lines and blue dashed lines, respectively.



Figure 5. A cladogram of family B1 GPCRs obtained following a maximum parsimonious analysis (1000 bootstrap replicates). The taxa are labelled using GenBank accession numbers and the species names. *Drosophila melanogaster* metabotropic glutamate receptor is utilized as an outgroup. Black symbols are used to denote sequences from *Homo sapiens* and colored symbols denote insect sequences. Note that the three receptor subtypes (CRF, PDF and CT/DH) form monophyletic clades. doi: 10.1371/journal.pone.0082466.g005

CT/DH-R1-C or *Rhopr-CT/DH-R2-B* is expressed in the dorsal vessel, salivary glands, hindgut and MTs which have previously

been shown to be targets of Rhopr-CT/DH [5,7,9]. Rhopr-CT/DH causes, at most, a 17-fold increase in the rate of fluid



Figure 6. Spatial expression analysis of *Rhopr-CT/DH-Rs* in fifth instar *R. prolixus* determined using quantitative PCR.

Rhopr-CT/DH-R1 (A) and *Rhopr-CT/DH-R2* (B) expression profile. Expression was analyzed in the following tissues: CNS (central nervous system), DV (dorsal vessel), Pool (fat bodies, abdominal nerves, diaphragms and trachea), FG (foregut), SG (salivary glands), AMG (anterior midgut), PMG (posterior midgut), MTs (Malpighian tubules), HG (hindgut), TST (testes), OV (ovaries) and PG (prothoracic glands). Expression of each variant for both the receptors is shown relative to *Rhopr-CT/DH-R1-B* transcript levels in salivary glands cDNA.

doi: 10.1371/journal.pone.0082466.g006



Figure 7. Spatial expression analysis of *Rhopr-CT/DH-Rs* in *R. prolixus* adult reproductive tissues determined using quantitative PCR.

Rhopr-CT/DH-R1 (A) and *Rhopr-CT/DH-R2* (B) expression profile. Expression was analyzed in the following tissues: TST (testes), M.R. (rest of the male reproductive tissues), OV (ovaries) and F.R. (rest of the female reproductive tissues). Expression of each variant for both the receptors is shown relative to *Rhopr-CT/DH-R1-B* transcript levels in testes cDNA. doi: 10.1371/journal.pone.0082466.g007

secretion by MTs compared to saline controls [5]; however, this rate is only 1.5% of the maximum rate stimulated by serotonin and Rhopr-CRF/DH. Perhaps Rhopr-CT/DH is used by the insect at times when diuresis needs to be maintained at a low level, such as the period after the rapid post-feeding diuresis and during digestion. Based on the expression patterns, Rhopr-CT/DH-R2 appears to be the one responsible for MT secretion in *R. prolixus*, whereas CT/DH-R1 is responsible for the diuretic function in D. melanogaster and A. aegypti [12,15]. FlyAtlas (www.flyatlas.org) data indicates that Drome-CT/DH-R2 is not expressed in D. melanogaster MTs. This observation suggests that the specialized functions, diuresis mediated through CT/DH-R2 in R.prolixus and through CT/DH-R1 in dipterans, have evolved independently. The phylogeny of these two receptor clusters has been discussed previously [31]. It remains to be seen whether CT/DH-R2 mediates diuresis in other non-dipteran species. Rhopr-CT/DH also increases muscle contractions of dorsal vessel, salivary glands and hindgut, all of which indicate a role for Rhopr-CT/DH in feedingrelated physiological events [7,9]. For instance, the contraction of salivary glands would aid in the release of saliva at the time of feeding. An increase in dorsal vessel contractions would result in increased circulation of the haemolymph, as well as the diuretic hormones that are present in the haemolymph following feeding [32]. Hindgut contractions aid in expulsion of waste, reduce unstirred layers around MTs and also increase haemolymph circulation. Hence increased contractility of dorsal vessel and hindgut could indirectly aid in post-feeding diuresis.

All the receptor transcripts are expressed with in the CNS but it is unclear which neural circuits Rhopr-CT/DH-Rs may be involved with. The gPCR analysis also identified novel target tissues of Rhopr-CT/DH; these include testes, ovaries and prothoracic glands. Rhopr-CT/DH, thus, may regulate reproductive physiology and ecydsteroidogenesis. CT/DH-like immunoreactivity is not associated with male or female reproductive tissues of R. prolixus (Zandawala and Orchard, unpublished). Hence the effect, if any, of Rhopr-CT/DH on reproductive tissues will most-likely be mediated via a hormonal route. Drome-CT/DH-R2 (hector) is expressed in a subset of *fruitless* neurons and has been shown to be critical for male courtship [33]. Consistent with this role, its transcript is enriched in *D. melanogaster* brain and male accessory glands [34]. It remains to be examined if Rhopr-CT/DH-R2 is also involved in courtship behaviour considering its expression in the CNS and reproductive tissues.

This is the first study to reveal that insects possess at least two CT/DH receptors, which may be functionally different. Our expression analysis suggests that Rhopr-CT/DH-Rs may mediate feeding-related physiological events, some of which must await further investigation. Moreover, we also identified novel target tissues for this neuropeptide, including testes, ovaries and prothoracic glands, suggesting a possible role for Rhopr-CT/DH in reproductive physiology and development.

Supporting Information

File S1. Tables S1-S6 and Figures S1-S3. Table S1: Primers used to amplify the partial cDNA sequence for Rhopr-CT/DH-R1 and Rhopr-CT/DH-R2. Table S2: Primers used to perform 5' RACE PCR reactions. Table S3: Primers used to perform 3' RACE PCR reactions. Table S4: Primers used to amplify the largest cDNA fragments. Table S5: Primers used to amplify full ORF and introduce Kozak sequence. Table S6: Primers used for qPCR reactions. Figure S1: Rhopr-CT/DH-R1-A cDNA sequence and the deduced amino acid sequence. The numbering for each sequence is shown at right. Within the nucleotide sequence, the exon-exon boundaries are shaded in gray and the potential polyadenylation signal is doubleunderlined. Within the amino acid sequence, the initial methionine start codon has been capitalized, the six conserved cysteine residues are shaded in red, the potential N-linked glycosylation sites are boxed and the predicted transmembrane domain is underlined. Figure S2: Rhopr-CT/DH-R2-A cDNA sequence and the deduced amino acid sequence. The numbering for each sequence is shown at right. Within the nucleotide sequence, the exon-exon boundaries are shaded in gray. Within the amino acid sequence, the initial methionine start codon has been capitalized, the conserved cysteine residues are shaded in red and the potential N-linked glycosylation sites are boxed. Figure S3: Kinetics of the bioluminescence responses of HEK/CNG (A) and CHO/G16 (B) cells expressing Rhopr-CT/DH-R1-B. Bioluminescence was recorded for every 5 seconds for 15 seconds following the addition of phosphate-buffered saline (PBS) or 10⁻⁶M peptide. Vertical bars represent SEM (n=3). Rhopr-CT/DH produced a rapid response, with the peak response for HEK/CNG cells and CHO/G16 cells between 5-10 seconds and 0-5 seconds, respectively. The assay was performed using the methods described earlier.

(DOCX)

Acknowledgements

The authors wish to thank Nikki Sarkar for maintaining the colony and Laura Sedra for dissecting adult reproductive tissues for qPCR analysis.

Author Contributions

Conceived and designed the experiments: MZ SL FH CJPG IO. Performed the experiments: MZ. Analyzed the data: MZ. Contributed reagents/materials/analysis tools: FH CJPG IO. Wrote the manuscript: MZ SL FH CJPG IO.

References

- Furuya K, Milchak RJ, Schegg KM, Zhang J, Tobe SS et al. (2000) Cockroach diuretic hormones: characterization of a calcitonin-like peptide in insects. Proc Natl Acad Sci U S A 97: 6469-6474. doi: 10.1073/pnas.97.12.6469. PubMed: 10841553.
- Coast GM, Garside CS, Webster SG, Schegg KM, Schooley DA (2005) Mosquito natriuretic peptide identified as a calcitonin-like diuretic hormone in *Anopheles gambiae* (Giles). J Exp Biol 208: 3281-3291. doi:10.1242/jeb.01760. PubMed: 16109890.
- Coast GM, Webster SG, Schegg KM, Tobe SS, Schooley DA (2001) The *Drosophila melanogaster* homologue of an insect calcitonin-like diuretic peptide stimulates V-ATPase activity in fruit fly Malpighian tubules. J Exp Biol 204: 1795-1804. PubMed: 11316500.
- Coast GM, Nachman RJ, Lopez J (2011) The control of Malpighian tubule secretion in a predacious hemipteran insect, the spined soldier bug *Podisus maculiventris* (Heteroptera, Pentatomidae). Peptides 32: 493-499. doi:10.1016/j.peptides.2010.11.009. PubMed: 21093508.
- Te Brugge VA, Lombardi VC, Schooley DA, Orchard I (2005) Presence and activity of a Dippu-DH31-like peptide in the blood-feeding bug, *Rhodnius prolixus*. Peptides 26: 29-42. doi:10.1016/j.peptides. 2004.08.025. PubMed: 15626502.
- Te Brugge V, Ianowski JP, Orchard I (2009) Biological activity of diuretic factors on the anterior midgut of the blood-feeding bug, *Rhodnius prolixus*. Gen Comp Endocrinol 162: 105-112. doi:10.1016/ j.ygcen.2009.01.025. PubMed: 19408362.
- Brugge VA, Schooley DA, Orchard I (2008) Amino acid sequence and biological activity of a calcitonin-like diuretic hormone (DH31) from *Rhodnius prolixus*. J Exp Biol 211: 382-390. doi:10.1242/jeb.013771. PubMed: 18203994.
- LaJeunesse DR, Johnson B, Presnell JS, Catignas KK, Zapotoczny G (2010) Peristalsis in the junction region of the *Drosophila* larval midgut is modulated by DH31 expressing enteroendocrine cells. BMC Physiol 10: 14. doi:10.1186/1472-6793-10-14. PubMed: 20698983.
- Orchard I (2009) Peptides and serotonin control feeding-related events in *Rhodnius prolixus*. Front Biosci (Elite Ed) 1: 250-262.
- Kaskani C, Poulos CP, Goldsworthy GJ (2012) The effects of linear and cyclic analogs of Locmi-DH, Dippu-DH(46) and Dippu-DH(31) on appetitive behavior in *Locusta migratoria*. Peptides 34: 258-261. doi: 10.1016/j.peptides.2011.04.009. PubMed: 21524674.
- Hewes RS, Taghert PH (2001) Neuropeptides and neuropeptide receptors in the *Drosophila melanogaster* genome. Genome Res 11: 1126-1142. doi:10.1101/gr.169901. PubMed: 11381038.
- Johnson EC, Shafer OT, Trigg JS, Park J, Schooley DA et al. (2005) A novel diuretic hormone receptor in *Drosophila*: evidence for conservation of CGRP signaling. J Exp Biol 208: 1239-1246. doi: 10.1242/jeb.01529. PubMed: 15781884.
- Evans BN, Rosenblatt MI, Mnayer LO, Oliver KR, Dickerson IM (2000) CGRP-RCP, a novel protein required for signal transduction at calcitonin gene-related peptide and adrenomedullin receptors. J Biol Chem 275: 31438-31443. doi:10.1074/jbc.M005604200. PubMed: 10903324.
- McLatchie LM, Fraser NJ, Main MJ, Wise A, Brown J et al. (1998) RAMPs regulate the transport and ligand specificity of the calcitoninreceptor-like receptor. Nature 393: 333-339. doi:10.1038/30666. PubMed: 9620797.
- Kwon H, Lu HL, Longnecker MT, Pietrantonio PV (2012) Role in diuresis of a calcitonin receptor (GPRCAL1) expressed in a distalproximal gradient in renal organs of the mosquito *Aedes aegypti* (L.). PLOS ONE 7: e50374.
- Kwon H, Pietrantonio PV (2013) Calcitonin receptor 1 (AedaeGPCRCAL1) hindgut expression and direct role in myotropic action in females of the mosquito *Aedes aegypti* (L.). Insect biochemistry and molecular biology 43: 588-593.
- Zandawala M, Lytvyn Y, Taiakina D, Orchard I (2012) Cloning of the cDNA, localization, and physiological effects of FGLamide-related allatostatins in the blood-gorging bug, *Rhodnius prolixus*. Insect

Biochem Mol Biol 42: 10-21. doi:10.1016/j.ibmb.2011.10.002. PubMed: 22061445.

- Reese MG, Eeckman FH, Kulp D, Haussler D (1997) Improved splice site detection in Genie. J Comput Biol 4: 311-323. doi:10.1089/cmb. 1997.4.311. PubMed: 9278062.
- Blom N, Gammeltoft S, Brunak S (1999) Sequence and structurebased prediction of eukaryotic protein phosphorylation sites. J Mol Biol 294: 1351-1362. doi:10.1006/jmbi.1999.3310. PubMed: 10600390.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731-2739. doi:10.1093/molbev/msr121. PubMed: 21546353.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948. doi:10.1093/bioinformatics/btm404. PubMed: 17846036.
- Visegrády A, Boros A, Némethy Z, Kiss B, Keseru GM (2007) Application of the BD ACTOne technology for the high-throughput screening of Gs-coupled receptor antagonists. J Biomol Screen 12: 1068-1073. doi:10.1177/1087057107309035. PubMed: 18087071.
- Seck T, Baron R, Horne WC (2003) The alternatively spliced deltae13 transcript of the rabbit calcitonin receptor dimerizes with the C1a isoform and inhibits its surface expression. J Biol Chem 278: 23085-23093. doi:10.1074/jbc.M211280200. PubMed: 12686555.
- Nag K, Sultana N, Kato Á, Hirose S (2007) Headless splice variant acting as dominant negative calcitonin receptor. Biochem Biophys Res Commun 362: 1037-1043. doi:10.1016/j.bbrc.2007.08.107. PubMed: 17822673.
- Stables J, Green A, Marshall F, Fraser N, Knight E et al. (1997) A bioluminescent assay for agonist activity at potentially any G-proteincoupled receptor. Anal Biochem 252: 115-126. doi:10.1006/abio. 1997.2308. PubMed: 9324949.
- Collin C, Hauser F, de Valdivia EG, Li S, Reisenberger J et al. (2013) Two types of muscarinic acetylcholine receptors in *Drosophila* and other arthropods. Cell Mol Life Sci 70: 3231-3242. doi:10.1007/ s00018-013-1334-0. PubMed: 23604020.
- Staubli F, Jorgensen TJ, Cazzamali G, Williamson M, Lenz C et al. (2002) Molecular identification of the insect adipokinetic hormone receptors. Proc Natl Acad Sci U S A 99: 3446-3451. doi:10.1073/pnas. 052556499. PubMed: 11904407.
- Conner AC, Simms J, Hay DL, Mahmoud K, Howitt SG et al. (2004) Heterodimers and family-B GPCRs: RAMPs, CGRP and adrenomedullin. Biochem Soc Trans 32: 843-846. doi:10.1042/ BST0320843. PubMed: 15494030.
- Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF et al. (2013) Molecular signatures of G-protein-coupled receptors. Nature 494: 185-194. doi:10.1038/nature11896. PubMed: 23407534.
- Mikhailova MV, Mayeux PR, Jurkevich A, Kuenzel WJ, Madison F et al. (2007) Heterooligomerization between vasotocin and corticotropinreleasing hormone (CRH) receptors augments CRH-stimulated 3',5'cyclic adenosine monophosphate production. Mol Endocrinol 21: 2178-2188. doi:10.1210/me.2007-0160. PubMed: 17536010.
- Park Y (2012) Endocrine regulation of insect diuresis in the early postgenomic era. Can J Zool 90: 507-520.
- Orchard I (1989) Serotonergic neurohaemal tissue in *Rhodnius* prolixus: Synthesis, release and uptake of serotonin. J Insect Physiol 35: 943-947. doi:10.1016/0022-1910(89)90017-6.
- Li Y, Hoxha V, Lama C, Dinh BH, Vo CN et al. (2011) The hector Gprotein coupled receptor is required in a subset of *fruitless* neurons for male courtship behavior. PLOS ONE 6: e28269. doi:10.1371/ journal.pone.0028269. PubMed: 22140564.
- Chintapalli VR, Wang J, Dow JA (2007) Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat Genet 39: 715-720. doi:10.1038/ng2049. PubMed: 17534367.